homo sapien homo sapien

QQQ1484 PQC3388 PQC3388 PQC1955 PC1955 PC1955 PC13258 PC13258

391 3.0 3924 390 3.0 2944 389.5 2.9 3866	380 2.9 1570 1 380 2.9 3969 1	379 2.9 1685 1 378 2.9 1183 1 375.5 2.8 3726 1 373 2.8 1763 1	44 368 2.8 1185 1 DRPL HUWAN 45 364 2.8 1426 1 BCL9_HUWAN 46 360 2.7 1509 1 GSR1 HUWAN 47 358 2.7 1618 1 NEST_HUWAN	355.5 2.7 2842 1 355 2.7 2090 1 353 2.7 1453 1	353 2.7 1983 1 352 2.7 1838 1 350 5 2 7 1736 1	350.5 2.7 2843 1 349.5 2.6 1822 1	348.5 2.6 1464 1 348 2.6 1669 1	348 2.6 346 2.6 344 2.6 341.5 2.6	341 2.6 1453 1 341 2.6 1815 1	340 2.6 1466 1 339.5 2.6 1341 1 339.5 2.6 1690 1 339 2.6 1960 1	338.5 2.6 1505 1 337.5 2.6 2426 1	335 2.5 1516 1 334 2.5 1516 1	332.5 2.5 332.5 2.5	332 2.5 810 1 330.5 2.5 1418 1	330.5 2.5 2349 1 330 2.5 1367 1	329.5 2.5 915 1 220 5 2.5 915 1	325 2.5	324 2.5 3149 1	323 2.4 1446 1	322.5 2.4 2004 1 320.5 2.4 1669 1	319.5 2.4 1362 1 319 2.4 1736 1	317 2.4 555 1	315.5 2.4 1459 1	315.5 2.4 1496 1 315 2.4 6632 1	314 2.4 1255 1 312.5 2.4 1712 1	311 2.4 2688 1 310.5 2.3 3358 1	309 2.3 1087 1	ATTACAMENT TE	ALIGNMENT	RESULT 1	
5.1.6 Compugen Ltd.		Search time 33 Seconds (without alignments) 3971.533 Million cell updates/sec	WDEBPKPLLCSQYETLSDSE 2517			ers: 141681				results predicted by chance to have a to the score of the result being printed, of the total score distribution.		Description	Q9y618 h nuclear r		O/53/6 homo Bapien O/99wubs rattus norv		Quing despite the party of the		Qojkse rattus norv		Q8nez4 homo sapien P34333 caenorhabdi	Q9qyx7 mus musculu O9wx48 raffus norv		V14086 II INCLEAR I P14873 mus musculu			ratt	omor Pomor	_	Q11428 homo sapien Q9xsj7 canis famil	
GenCore version 5 Copyright (c) 1993 - 2004 Cc	OM protein - protein search, using sw model	Run on: April 15, 2004, 19:19:09 ; Sea (wi 397	Title: US-09-522-753-5 Perfect score: 13215 Sequence: 1 MSGSTQLVAQTWRATEPRYP	Scoring table: BLOSUM62 Gapop 10.0 , Gapext 0.5	Searched: 141681 segs, 52070155 residues	er of	Minimum DB ség length: 0 Maximum DB seg length: 200000000	Post-processing: Minimum Match 0% Maximum Match 100% Listing first 100 summaries	Database : SwissProt_42:*	Pred. No. is the number of results predicted score of score greater than or equal to the score of and is derived by analysis of the total scor	SUMMARIES	Result Query No. Score Match Length DB ID	13215 100.0 2517 1	10987.5 83.1 2472 1 4187 31.7 2453 1		569.5 4.3 3664 I 575 4 0 2142 1	9 503 3.8 2716 0 492 5 3.7 2774	1 470.5 3.6 5262 1	3 462.5 3.5 5085 1	4 451.5 3.4 2805 1 5 439.5 3.3 2067 1	6 439 3.3 4911 1 7 438.5 3.3 1780 1	8 434.5 3.3 5038 1 9 432 3.3 2167 1	0 428.5 3.2 2715 1	2 424.5 3.2 2464 1	3 415.5 3.1 1226 1 4 410.5 3.1 5147 1	5 410 3.1 2468 1 6 407 3.1 5120 1	7 406.5 3.1 2459 1	399 3.0 2161	1 396 3.0 4903 1 1 396 3.0 3703 1	3 391.5 3.0 1411 1 3 391.5 3.0 1460 1	

mus musculu mus musculu mus musculu mus musculu mus musculu rattus musculu rattus norv homo sapien dus musculu mus musculu homo sapien mus musculu mus

QG1315 P12034 P12034 P12034 P0911u4 P0

LIGNMENTS

```
INTEGRATION TO THE WINDOWN THE WAS ASSENDED TO THE WAS ASSENDED TO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TISSUE=Fetal liver;
MEDLINE=96408715; PubMed=8813722;
Sande S., Privalsky M.L.;
Sande S., Privalsky M.L.;
"Identification of TRACs (T3 receptor-associating cofactors), a family
of cofactors that associate with, and modulate the activity of,
                                                                                                                                                                                                                                                                                                                                                                                                    TIŜSUE=Pituitary;
MEDIINE=99178941; PubMed=10077563;
Ordentlich P., Downes M., Xie W., Genin A., Spinner N.B., Evans R.M.;
Ordentlich P., Downes M., Xie W., Genin A., Spinner N.B., Evans R.M.;
"Unique forms of human and mouse nuclear receptor corepressor SWRT.";
Proc. Natl. Acad. Sci. U.S.A. 96:2639-2644(1999).
                            MCR2 HUMAN STANDARD; PRT; 2517 AA.

09Y618; 000613; 015416; 013354; 09Y5U0;
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2003 (Rel. 42, Last annotation update)
Nuclear receptor co-repressor 2 (N-CoR2) (Silencing mediator of retinoic acid and thyroid hormone receptor) (SMRT) (SMRTe) (Thyroid-retinoic-acid-receptor-associated co-repressor) (SMRT) (Thyroid-normone receptor)
Retinoic-acid-receptor-associated co-repressor) (T3 receptor-associating factor) (TGC repeat protein 26).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE OF 1023-2517 FROM N.A.
TISSUE=Cervical adenocarcinoma;
TISSUE=Cervical adenocarcinoma;
TISSUE=Cervical adenocarcinoma;
Chen J.D., Evans R.M.;
"A transcriptional co-repressor that interacts with nuclear hormone
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TISSUE=Brain cortex;
MEDLINE=97369492; PubMed=9225980;
Margolis R.L., Abraham M.R., Gatchell S.B., Li S.-H., Kidwai A.S.,
Margolis R.L., Abraham M.R., Catchell S.B., Li S.-H., Kidwai A.S.,
Margolis R.L., Abraham M.C., Callahan C., McInnis M.G., Ross C.A.;
Breschel T.S., Stine O.C., Callahan C., McInnis M.G., Ross C.A.;
"CDNAB with long CAG trinucleotide repeats from human brain.";
Hum. Genet. 100:114-122(1997).
                                                                                                                                                                                                                                                                 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         INTERACTION WITH MINT.

WEDLINE=21231190; PubMed=11331609;

Shi Y., Downes M., Xie W., Kao H.-Y., Ordentlich P., Tsai C.-C.,
Shi Y., Downes M.,

Hon M., Evans R.M.;

"Sharp, an inducible cofactor that integrates nuclear receptor

"Sharp, an inducible cofactor that integrates nuclear receptor

repression and activation.";

Genes Dev. 15:1140-1151(2001).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TISSUB=Cervical adenocarcinoma; MEDLINE=99199215; PubMed=10097068; Park E.J., Schroen D.J., Yang M., Li H., Li L., Chen J.D.; SMTFe, a silencing mediator for retinoid and thyroid hormone receptors-extended isoform that is more related to the nuclear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       receptor corepressor.";
Proc. Natl. Acad. Sci. U.S.A. 96:3519-3524 (1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A. (ISOFORM TRAC-1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Mol. Endocrinol. 10:813-825(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A. (ISOFORM SMRT).
                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A. (ISOFORM SMRT).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE OF 428-613 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             hormone receptors
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Nature 377:454-457(1995)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         INTERACTION WITH HDAC10.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          receptors."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               nuclear
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce) or send an email to license@ib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Isold=09Y618-2; Sequence=VSP 003412, VSP 003413;
and acts as an antirepressor.
InDUCTION: Regulated during cell cycle progression are detected in lung, spleen and brain.
InDUCTION: Regulated during cell cycle progression functions that are divided into three independent repression domains (RD). RD2 and RD3). The C-terminal region contains the nuclear receptorand nature of the contains the nuclear receptorand comains (ID) and ID2.
Interacting domains that are divided in two separate interaction domains (ID) and ID2.
Interacting domains that of contains a conserved sequence referred to a the CORNR box. This motif is required and sufficient to permit binding to unliganded TR and RARS. Sequences of flanking the CORNR box determine nuclear hormone receptor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam; PF00249; myb DNA-binding; ...
Pfam; PF00249; myb DNA-binding; ...
PR03TTS; SS0090; MYB 3; 1.
PR05TE; PS50090; MYB 3; 1.
Nuclear protein; Transcription regulation; DNA-binding; Repressor;
Coiled coil; Alternative splicing; 3D-structure.
Coiled coil; Alternative splicing; S-structure.
Coiled coil; Alternative splicing; S-structure.
Coiled coil; Alternative splicing; S-structure.
Coiled coil.; Alternative splicing; S-structure.
Coiled coiled coil.; Alternative splicing; S-structure.
Coiled coiled coil.; Alternative splicing; S-structure.
Coiled c
nuclear receptors by promoting chromatin condensation, thus preventing access of the basal transcription.

SUBUNIT: Interacts with HDAC7 (By similarity). Forms a large corpersesor complex that contains SUBJA/B and histone deacetylases HDAC1 and HDAC2. This complex associates with the thyroid (TR) and the retinoid acid receptors (RAR) in the absence of ligand, and may stabilize their interacts with TFIIB. The SRMT isoform interacts with HDAC10. Interacts with MINT.

SUBCELLUIAR LOCATION: Nuclear.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             specificity.
-!- SIMILARITY: Contains 1 SANT-A domain.
-!- SIMILARITY: Contains 1 Myb-like domain.
-!- SIMILARITY: BELONGS TO THE N-COR NUCLEAR RECEPTOR COREPRESSORS
-!- SIMILARITY: BELONGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MIM; 600848; -.
Oc) GO:0005634; C:nucleus; TAS.
GO; GO:000374; F:transcription co-repressor activity; TAS.
GO; GO:0003714; F:transcription co-repressor activity; TAS.
InterPro; IPR01005; Myb_DNA_binding.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Event=Alternative splicing; Named isoforms=2;
Name=SMRT; Synonyms=TRAC-2;
IsoId=Q9Y618-1; Sequence=Displayed;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRO-RICH.
CORNR BOX OF ID1.
CORNR BOX OF ID2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POLY-GLN.
POLY-LYS.
POLY-PRO.
POLY-PRO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; AF113003; AAD20946.1; --
EMBL; AF125672; AAD22973.1; --
EMBL; U3146; AAD52036.1; --
EMBL; G83390; AAB50847.1; --
EMBL; U80750; AAB91446.1; --
PTR; 560255; S60255.
PDB; IKKQ; 27-FEB-02.
TRANSFAC; T04689; --
Genew; HGNC; 7673; NCOR2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           820
2143
2346
510
685
1389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
DNA_BIND
DNA_BIND
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
```

661 NIDRILOGHKI, KMERGENNERKKKKABANASEBAAPPRVEEDENEAGOVGKBEBNYEE 720 661 NIDBILOGHKI, KMERGENNERKKKKABANASEBAAPPRVEEDENEAGOVGKREBNYEE 720 721 AEALIANGGNEVPREGGGGPATVNNSSDTSSIESHTEAAKDTGONGPREPATLCADGEP 760 721 AEALIANGGNEVPREGGGGPATVNNSSDTSSIESHTEAAKDTGONGPREPATLCADGEP 760 721 AEALIANGGNEVPREGGGGPATVNNSSDTSSIESHTEAAKDTGONGPREPATLCADGEP 760 722 AEALIANGGNEVPREGGGGPATVNNSSDTSSIESHTEAAKDTGONGPREPATLCADGEP 760 723 AEALIANGGNEVPREGGGGPATVNNSSDTSSIESHTEAAKDTGONGPREPATLCADGEP 760 724 GGPOTPHERITSRALIEPTABEATGAPPPAPERSPERAKOTABANGTGONGPREGGG 840 841 BEGOKPPALEELAVDTGAAEBVCKTEERAEGGPAKGACABANATAGAAPPVEE 840 842 BEGOKPPALEELAVDTGAAEBVCKTEERAEGGPAKGACABANATAGAAPPVEE 840 843 BEGOKPPALEELAVDTGAAEBVCKTEERAEGGPAKGACABANATAGAAPPVEE 840 844 BEGOKPPALEELAVDTGAAEBVCKTEERAEGGPAKGACABANATAGAAPPVEE 840 844 BEGOKPPALEELAVDTGAAEBVCKTEERAEGGPAKGACABANATAGAAPPVEE 840 844 BEGOKPPALEELAVDTGAAEBVCKTEERAEGGPAKGACABANATAGAAPPVEE 840 845 BEGOKPPALEELAVDTGAAEBVCKTEERAEGGPAKGACABANATAGAAPPVEE 840 846 BEGOKPPALEELAVDTGAAEBPREDAAAPPVEE 840 847 BEGOKPPALEELAVDTGAAEBPREDAAGATGAAPPVEE 840 848 BEGOKPPALEELAVDTGAAEBPREDAAGATGAAPPVEE 840 849 BEGOKPPALEELAVDTGAAEBATGAAEBPREDAAFKAAPPPEAGAARATAGAAPPVEE 840 840 GGATTAAKSGAAPQASAAAAPPTAGAABAATGAAPPVEE 840 841 BEGOKPPALEELAVDTGAAEBATGAAEBATGAAPPPEAGAARATAGAAPPVEE 840 840 GGATTAAKSGAAAAAPPTAGAAAAPPTAGAABAAPTAGAABAAATAGAAPAAPPTAGAAAPPAAAAAPATGAAAPPAABAATGAAAPAAAAPA	1621 VDLYRSHIPLAFDPTSIPRGIPLDAAAYYLPRHIAPNPTYPHLYPPYLIRGYPDTAALE 1680 1621 VDLYRSHIPLAFDPTSIPRGIPLDAAAAYYLPRHIAPNPTYPHLYPPYLIRGYPDTAALE 1680 1681 NRQTIINDYITSQQMHINTATAMAQRADMLRGLSPRESSLALNYAAGPRGIIDLSQVPHL 1740
B	\$ 8 \$ \$ \$
Promobine 1842 1846 Poly-GLY.	91 DKEDLLKEKTDDTSGEDNDEKEAVASKGRKTANSGERKGEITRSMANEANSEERITFOX

```
2400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2460
                                                                                                                                                                                                                                                                                                                                                                                                                             2280
                                                                                                                                                                                                                                                                                                                                                                                                                                                      2340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2100
                                                                                                                                                                                                                                                              2160
                                                                                                                                                                                                                                                                                                                APLYSFPGASCPVLDLRRPPSDLYLPPPDHGAPARGSPHSEGGKRSPEPNKTSVLGGGED 2220
                                                                                                                                                                                                                                                                                                                                                                  PFPYNDLIMKLQAGVMASPPPGLPAGSGPLAGPHHAWDEBPKPLLCSQYETLSDSE 2517
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2461 PFPYNPLIMRLQAGVWASPPPPGLPAGSGPLAGPHHAMDEEPKPLLCSQYETLSDSB 2517
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A. (ISOFORMS ALPHA AND BETA).
TISSUB-Brain, and Spleen;
MEDLINE=99178941; PubMed=10077563;
Ordentlich P., Downes M., Xie W., Genin A., Spinner N.B., Evans R.M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GKAKVSGRPSSRKAKSPAPGLASGDRPSVSSVHSEGDCNRRTDLTNRVWEDRPSSAGST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCR2 MOUSE STANDARD; PRT; 2472 AA.
O9W143; O9WUC1;
16-OCT-2001 [Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
10-OCT-2003 (Rel. 40, Last sequence update)
10-OCT-2003 (Rel. 42, Last annotation update)
Nuclear receptor oc-repressor 2 (N-CoR2) (Silencing mediator of Nuclear receptor oc-repressor acid hormone receptor)
retinoic acid and thyroid hormone receptor (SWRT) (SWRT) (Thyroid-, retinoic-acid-receptor-associated co-repressor) (T3 receptor-retinoic-acid-receptor-associated co-repressor) (T3 receptor-
                                                                                                                                                                                                                                                                                                                                                                                                    GIEPVSPPEGMTEPGHSRSAVY PLLYRDGEQTEPSRMGSKSPGNTSQPPAFFSKLTESNS
                                                                                                                                                                                                                                                                                                                                                                                                                  AMVKSKKQBI NKKLNTHNRNEPBYNI SQPGTEI FNMPAITGTGLMTYRSQAVQEHASTNM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GLEAIIRKALMGKYDQWEESPPLSANAFNPLNASASLPAAMPITAADGRSDHTLTSPGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LGYHGSSYSPEGVEPVSPVSPSTTHDKGLPKHLEELDKSHLEGELRPKQPGPVKLGGEA
                                                          LGGTLDGVYPTLMEPVLLPKEAPRVARPERPRADTGHAFLAKPPARSGLEPASSPSKGSE
                                                                                                                                                               PRPLVPPVSGHATIARTPAKNLAPHHASPDPPAPASASDPHREKTQSKPFSIQELELRS
                                                                                                                                                                                                             AHLPHLRPLPESOPSSSPLLQTAPGVKGHORVVTLAQHISEVITQDYTRHHPQQLSAPLP
                                                                                          SPISPRTQDALQQRPSVLHNTGMKGIITAVEPSKPTVLRSTSTSSPVRPAATFPPATHCP
                                        DRERDRDREREKSILTSTTTVEHAPIWRPGTEQSSGSSGGGGGSSSRPASHAHQH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Mus musculus (Mouse).
Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Buteleostomi,
Mammalia, Eutheria, Rodentia, Sciurognathi, Muridae, Murinae, Mus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         associating factor) (TRAC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCOR2 OR SMRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2461
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2401
                                                                                                                                                                                                                                                                                                                                                                                                            2221
                                                                                                                                                                                                                                                                                                                                                                                                                                                             2281
                                                                                                                                                                                                                                                       2041
                                                                                                                                                                                                                                                                                2041
                                                                                                                                                                                                                                                                                                         2101
                                                                                                                                                                                                                                                                                                                                                          2161
                                                                         1801
                                                                                                                         1861
                                                                                                                                                   1921
                                                                                                                                                                            1921
                                                                                                                                                                                                     1981
                                                                                                                                                                                                                              1981
                                                1801
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCR2_MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          В
                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                   원
                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                            셤
                                                                                                                                                                                                    ò
                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                         ద
                                                                                                                                                                                                                                                                                                    ઠે
                                                                                                                                                                                                                                                                                                                                                         ò
                                                                       원
                                                                                                  ò
                                                                                                                                                  ò
                                         ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE=20107033; PubMed=10640276;
Kao H.-Y., Downes M., Ordentlich P., Evans R.M.;
Kao H.-Y., Downes M., Ordentlich P., Evans R.M.;
Isolation of a novel histone deacetylase reveals that class I and class II deacetylases promote SMRT-mediated repression.";
Genes Dev. 14:55-66(2000).
--- FUNCTION: Mediates the transcriptional repression activity of some nuclear receptors by promoting chromatin condensation, thus preventing access of the basal transcription.
---- SUBUNIT: Forms a large corepressor complex that complex associates and histone deacetylases HADAC1 and HDAC2. This complex associates with the thyroid (TR) and the retinoid acid receptors (RAR) in the absence of ligand, and may stabilize their interaction with TFIIB.

Interacts with HDAC10 and MINT (By similarity). Interacts with
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             embryos.

DOMAIN: The N-terminal region contains repression functions that are divided into three independent repression domains (RDI, RD2 and RD3). The C-terminal region contains the nuclear receptor—interacting domains that are divided in two separate interaction domains (IDI and IDS).

DOMAIN: The two interaction domains (ID) contain a conserved bequence referred to as the CORNR box. This motif is required and sufficient to permit binding to unliganded TR and RARS. Sequences flanking the CORNR box determine nuclear hormone receptor
corepressor SMRT.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        There, prize to the state of th
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   specificity.
SIMILARITY: Contains 1 SANT-A domain.
SIMILARITY: Contains 1 Myb-like domain.
SIMILARITY: BELONGS TO THE N-COR NUCLEAR RECEPTOR COREPRESSORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Name=Beta;
Isold=Q9WU42-2; Sequence=VSP_001414;
TISSUE SPECIFICITY: Ubiquitous. Also widely expressed in early
                                                                                                                                                             MEDLINE=99199215; PubMed=10097068; Park B.J., Schroen J.D.; Park B.J., Schroen D.J., Yang M., Li H., Li L., Chen J.D.; SmRTe, a silencing mediator for retinoid and thyroid hormone "SMRTe, a silencing mediator for retinoid and thyroid hormone receptors-extended isoform that is more related to the nuclear
          forms of human and mouse nuclear receptor
tl. Acad. Sci. U.S.A. 96:2639-2644(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Event=Alternative splicing; Named isoforms=2;
                                                                                                                                                                                                                                                                                                                             Proc. Natl. Acad. Sci. U.S.A. 96:3519-3524(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Name=Alpha;
IsoId=Q9WU42-1; Sequence=Displayed;
                                                                                                      FROM N.A. (ISOFORM ALPHA).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SUBCELLULAR LOCATION: Nuclear. ALTERNATIVE PRODUCTS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; AF113001; AAD20944.1; -. EMBL; AF113002; AAD20945.1; -.
                    "Unique forms of human
Proc. Natl. Acad. Sci.
                                                                                                                                                                                                                                                                                                                                                                                 [3]
INTERACTION WITH HDAC7
                                                                                                                                                                                                                                                                                                                   corepressor.
                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -
```

16 1.5	1168
6 B 6 B 6 B 6 B 6 B 6 B 6 B 6 B 6 B 6 B	6 6 6 6 6 6
MYB. COLLED COLL (POTENTIAL). PRO-RICH. PRO-RICH. PRO-RICH. PRO-RICH. CORNR BOX OF ID1. CORNR BOX OF ID1. CORNR BOX OF ID2. POLY-GLN. POLY-ALA. M -> RL (IN REF. 1; AAD20944). T -> M (IN REF. 1; AAD20945). E -> K (IN REF. 2). FHAADPSA> TRADPL (IN REF. 2). FHAADPSA> TRADPL (IN REF. 2). FRASTING (IN REF. 2). GSATGGSITKGLESTRAADGPSYRGSITHG> A SPRASOTYPERADGPSYRGSITHG> A STRACT CREF. 2). MISSING (IN REF. 2). GSATGGSITKGLESTRAADGPSYRGSITHG> A STRACT CREF. 2). MISSING (IN REF. 2). GSATGGSITKGLESTRAADGPSYRGSITHG> A STRACT CREF. 2). MISSING (IN REF. 2). GSATGGSITKGLESTRAADGPSYRGSITHG> A STRACT CREF. 2). MISSING (IN REF. 2). G -> A (IN REF. 2). T -> A (IN REF.	GSEDLTKDRSLAGKLEPVSPPSPHADPELELAPSRLSKEELJQNMDRVD GSEDLTKDRSLAGKLEPVSPPSPHADPELELAPSRLSKEELJQNMDRVD SKLKKKQQQLEEEAAKPPEPEKPVSPPPIESKHRSLVQIIYDENRKKAEA 2 SKLKKKQQQLEEEAAKPPEPEKPVSPPPIESKHRSLVQIIYDENRKKAEA 2 QVELPLYNQPSDTRQYHENIKINQAMRKKLILYFKRRNHARKQWKQKFCQ 3 QVELPLYNQPSDTRQYHENIKINQAMRKKLILYFKRRNHARKQWEGRFCQ 3 CVELPLYNQPSDTRQYHENIKINQAMRKKLILYFKRRNHARKQWEGRFCQ 3 KKVERIENNPRRAKESKVREYYEKQPEIRKQRELQERMQSRVGQRGGG 3 KKVERIENNPRRAKESKVREYYEKQPEIRKQRELQERMQSRVGQRGGG 3
609 653 652 668 652 668 989 999 1351 1357 2294 2098 2296 2300 434 2437 36 254 36 254 36 254 36 254 36 254 36 1069 113 1133 1149 1157 1172 1201 1172 1201 1173 1133 1149 1157 1157 1201 1157 1201 1158 1857 1159 1959 1913 1149 1157 1201 1157 1201 1201 1201 1201 1201 1201 1201 1201	SPLLATGQPSG SPLLATGQPSG SEITWVEQOIS REITWVEQOIS AHRILEGLGPG AHRILEGLGPG AHRILEGLGPG AHRILEGLGPG AHRILEGLGPG AHRILEGLGPG AHRILEGLGPG AHRILEGLGPG AHRILEGLGPG
S M	
5 1 1 7 7 8 BEICHCHOOCH COOCH	
DNA BIND DOMAIN CONFLICT CONF	121 181 181 181 241 301

```
2483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2079
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2019
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1903
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2005
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2064
                                                                                                                                                                                                                   YYLPRHLAPNPTYPHLYPPYLLRGYPDTAALENRGTIINDYITSQQMHHNTATAWAQRAD 1708
                                                                                                                                                                                                                                                                        MLRGLSPRESSLALNYAAGPRGIIDLSQVPHLPVLVPPTPGTPATAMDRLAYLPTAPQPF 1768
                                                                                                                                                                                                                                                                                        AHEGLVATVKEAGRSIHEIPREEIRHTPELPLAPRPLKEGSITOGTPLKYDTGASTTGSK 1469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LPPDHGAPARGSPHSEGGKRSPEPNKTSVLGGGEDGIEPVSPPEGMTEPGHSRSAVYPL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NISQPGTEIFNMPAITGTGLMTYRSQAVQEHASTNMGLEAIIRKALMGKYDQWEESPPLS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ANAFNPLNASASLP-AAMPITAADGRSDHTLISPGGGGKAKVSGRPSSRKAKSPAPGLAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ANAFNPLNASASLPAAAMPITTADGRSDHALTSPGGGGKAKVSGRPSSRKAKSPAPGLAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GDRPPSVSSVHSEGDCNRRTPLTNRVWEDRPSSAGSTPFPYNPLIMRLQAGVMASPPPPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GVKGHQRVVTLAQHISEVITQDYTRHHPQQLSAPLPAPLYSFPGASCPVLDLRRPPSDLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2140 LPPPDHGTPARGSPHSEGGKRSPEPSKTSVLGSSEDAIEPVSPPEGMTEPGHARSTAYPL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LYRDGEQTEPSRMGSKSPGNTSQPPAFFSKLTESNSAMVKSKKQEINKKLNTHNRNEPEY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PGTEQSSGSSGSGGGSSSSRPASHSHAHOHSPISPRTODALOORPSVLHNTGMKGIIT
                                                                                                                                                                                                                                                                                                                                                                                                   1889 AVEPSKPTVL---RSTSTSSPVRPAATFPPATHCPLGGTLDGVYPTLMEPVLLPKEAPRV
                                                                                                                                                                                                                                                                                                                                                                                                                                                      ARPERPRADIGHAFLAKPPARSGLEPASSPSKGSEPRPLVPPVSGHATIARTPAKNLAPH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ARPERARVDAGHAFLTRPPGR---EPASSPSKSSEPRSLAPPSSSHTALARTPAKNLAPH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HASPDPPAPPASASDPHREKTQSKPFSIQELELRSLGYH-GSSYSPEGVEPVSPVSSPSL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 THDKGLPKHLEELDKSHLEGELRPKQPGPVKLGGERAHLPHLRPLPESQPSSSPLLQTAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1500 VPELGKPRQSPLTYEDHGAPFTSHLPRGSPVTTREPTPRLQEGSLLSSKASQDRKLTSTP
                                                                                                                                                              REIAKSPHSTVPEHHPHPISPYEHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLD-AAAA
                                                                                                                                                                                                                                   SSRHSSSPLSPGGPTHLTKPTTTSSSERERDRDRERDREREKSILTSTTVEHAPIWR
               VPELGKPROSPLTYEDHGAPFAGHLPRGSPVTWREPTPRLOEGSLSSSKASQDRKLTSTP
                                                     KHDVRSLIGSPGRTFPPVHPLDVMADARALERACYEESLKSRPGTASSSGGSIARGAPVI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2259
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1946
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1904
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2006
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2065
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2020
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2080
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2245
                                                                                                                                                                                                                                                                                                                                                                                          1829
                                                                                                                                                                                                                                                                                                                                                                                                                     1794
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1961
                                                                                                                1530
                                                                                                                                                                      1590
                                                                                                                                                                                                1560
                                                                                                                                                                                                                            1649
                                                                                                                                                                                                                                                                                1709
        1410
                                                            1470
                                                                                                                                                                                                                                                                                                                                      1769
```

位 昭 位 田 位 田 位 田 位 田 位 田 位 田

ð

상 . a

셤

g

ð

8 8

6 6 6 6 6

8

8 8 8 8

```
DOMAIN: The V-terminal region contains repression functions that are divided into three independant repression domains (RDI, RD2 and RD3). The C-terminal region contains the nuclear receptorinteracting domains that are divided in two separate interaction domains (IDI and ID2).

DOMAIN: The two interaction domains (ID) contain a conserved bowain: The two interaction domains (ID) contain a conserved sufficient to permit binding to unliganded TR and RARS. Sequences flanking the CORNR box determine nuclear hormone receptor
                                                                                                                                                                                                                                                                                                                                                                                                                            TISSUE-Liver;
MEDLINE=95280959; PubMed=7760852;
Seol W., Choi H.S., Moore D.D.;
Isolation of proteins that interact specifically with the retinoid "Isolation for proteins that interact specifically with the retinoid receptor: two novel orphan receptors.";
Mol. Endocrinol, 9:72-85(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COREPRESSORS
                                                                                                                                                                                                            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Butheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                             "Ligand-independent repression by the thyroid hormone receptor mediated by a nuclear receptor co-repressor."; Nature 377:397-404(1995).
                                                                                               NCRI_MOUSE STANDARD; PRT; 2453 AA.

Q60974; Q60812;

16-0CT-2001 (Rel. 40, Last sequence update)

16-0CT-2003 (Rel. 42, Last annotation update)

10-0CT-2003 (Rel. 42, Last annotation update)

Nuclear receptor co-repressor 1 (N-CoR1) (N-CoR) (Retinoid X interacting protein 13) (RIP13).

NCORI OR RXRIP13.
                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A. .....TISSUE=Pituitary;
TISSUE=Pituitary;
MEDLINE=96008339; PubMed=7566114;
MEDLINE=96008339; Naeaer A.M., Heinzel T., Torchia J., Gloss B., Hoerlein A.J., Naeaer A.M., Heinzel T., Torchia J., Gloss B.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RECEPTOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Event=Alternative splicing; Named isoforms=2;
                           2472
                                                                                                                                                                                                                                                                                                                                                                                                    [2] SEQUENCE OF 1792-2453 FROM N.A. (ISOFORM LONG)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  specificity.
SIMILARITY: Contains 1 SANT-A domain.
SIMILARITY: Contains 1 Myb-like domain.
SIMILARITY: BELONGS TO THE N-COR NUCLEAR
Ā
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Isold=060974-2; Sequence=VSP_003411; TISSUE SPECIFICITY: Ubiquitous.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      IsoId=Q60974-1; Sequence=Displayed;
                                                                                                                                                                                                          Mus musculus (Mouse)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ++
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     + + +
                                                                                                  à
                         셤
```

```
SRSPAPPADKEAFAAEAQKLPGDPPCWTSGLPFPVPPREVIKASPHAPDPSAFSYAPPGH 1082
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RLP-----TTRPTRPPPPLIPSSKTTVASEK-PSFI--MGGSISOGTPGTYLSSHNQAYP 1086
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPA------DVLYKGTITRI-IG 1248
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PIGTPVSGYALYORHIKAMHESALLE-------EQRQRQEQVDLECRSSTSPCST 1008
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1083 PLPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISQG-----MSVQLHVPY 1137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEASVLRG 1197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETAAPK 1308
    DRQFMNVWTDHEKEIFKDKFIQHPKNFGLIASYLERKSVPDCVLYYYLTKKNENYKALVR 492
                                                                                      YFNYKKRONLDBILQQHKLKMEKERNARRKKKKAPAAASEEAAFPPVVEDEEMEASGVSG 712
                                                                                                                                                                                                                                                                                                                                                         EGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSSGAPQ---DSDSSATCSADEVDE 928
                                                                                                                                                                                                                                                                                                                                                                                                                     :
EPPPPLPPPPEPISTEPVETSRWTEEEMEVAKKGLVEHGRNWAAIAKNVGTKSEAQCKNF
                                                                                                                                                                                                                              NEEEMVEEAEALHASGNEVPRGECSGPATVNNSSDTESIPSPH-TEAAK---DTGQNGPK
                                                                                                                                                                                                                                                             PPATLGADGPPPGPPTPPRRTSRAPIEPTPASEATGAPTPPPAPPSPSAPPDVVPKEEKE
                                                                                                                                                                                                                                                                                                                                  795 VEAQVTDSASAETAEPMDVDHEECGAEGSSVLDPPAPTK--ADSVDPEMQV-PENTASKG
                                                                                                                                                                                                                                                                                                                                                                                EGDAKERDLEST-----SEKTEARDEDVVVAEQIERPEPQSDDDSSATCSADEGVD
                          RSY-RRRGKSQQQQQQQQQQQQQQQQPMPRSSQEEXDEKEKEKEAEK--EEEKPEVEND
                                                                       KEDLLKEKTDDTSGEDNDEKEAVASKGRKTANSOGRRKGRITRSMANE---ANSEEAITP
                                                                                                                                                                                                                                                                                    ------SSAVPITKPAERES
                                                                                                                                                                                                                                                                                                            EE-----TAAAPPVEEGEEQ-----KPPAAEELAVDTGKAEEPVKSECTEEAE
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ------OVTKVHEPPREDAAPTKPAPPAPPPQNLQPESDAPQQPGSSPRGK
                                                                                                                      QQSAEL-----ASMELNESSRWTEEEMETAKKGLLEHGRNWSAIARMVGSKTVSQCKNF
                                                                                                                                                                                                                                                                                                                                                                                                       AEGGDKNRL--LSPRPSLLTPTGDPRANASPQKP--LDLKQLKQRAAAIPPI-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | | | :: | | :: | | | | | | | | | : | | ESSPEKV---REGARSPRIAHEMSL-K
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RTYDMMEGRVGRAIS ----SASIEGLMGRAIPPERHSPH-HLKEQHHIRGSITQGIPR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SYVEAQEDYLRREAKLLKREGTPPPPPPPSRDLTEAYKTQALGPLKLKPAHEGLVATVKEA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GRSIHEIPREEL----RHTPELPLAPRPLKEGSITQGTPLKYDTGASTTGSKKHDVRSL
                                                                                                                                                                                          YFNYKRRHNLDNLLQQHKQKASRKPREERDVSQCESVASTVSA----QEDEDIEAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1035 1
                                                                                             545
                                                                                                                                            605
                                                                                                                                                                                                                                                               169
                                                                                                                                                                                                                                                                                                                                                                                                       929
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1146
   433
                          485
                                                  493
                                                                      542
                                                                                                                     599
                                                                                                                                                                   653
                                                                                                                                                                                          665
                                                                                                                                                                                                                713
                                                                                                                                                                                                                                      717
                                                                                                                                                                                                                                                                                                            829
                                                                                                                                                                                                                                                                                                                                                                                852
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1198
                                                                                                                                                                                                                                                                                                                                                         872
                                                                                                                                                                                                                                                                                                                                                                                                                            903
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1023
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1309
                                                                                                                                                                                                                                                                                    757
                                                                                                                                                                                                                                                                                                                                                                                                                                                   977
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1359
                                           g
                                                                                         셤
                                                                                                                                       셤
                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                      ઠે
                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                       g
                       ઠે
                                                                Š
                                                                                                                     Š
                                                                                                                                                                 ò
                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                                                                                                                                                       ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   .iq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         음
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the BMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PGNERSQELHLRPESHSYLPELGKSEMEFIESKRPRLELLPDPLLR-----PSPLLATGQ 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DRQVMNWSEQEKETFREKFWQHPKNFGLIASFLERKTVAECVLYYYLTKKNENYKSLVR 484
                                                                 LMEALEKKVERIENNPRRRAKESKVREYYEKOFPEIRKORELOERMOSRVGORGSGLSMS 364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AARSEHEVSEIIDGLSEQENLEKQMRQLAVIPPMLYDADQQRIKFINMNGLMADPMKVYK 424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   KVEQQILKLKKKQQQLEEEAAKPPEPEKPVSPPPVEQKHRSIVQIIYDENRKKAEEAHKI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                254 PEGLGPKVELPLYNQPSDTKVYHENIKTNQVMRKKLILFFKRRNHARKQREQKICQRYDQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EPRYPPHSLSYPVQIARTHTDVGLLEYQ--HHSRDYASHLSPGSIIQPQRRRPSLLSEFQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 OSRYPSHSVQYTFPSARHQQEFAVPDYRSSHLEVSQASQLLQQQQQQQQLRRRPSLLSEFH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PAG---SEDLTKDRSLTGKLE-PVSPPSPPHTDPELELVPPRLSKEELIQNMDRVDREIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MVEQQISKLKKKQQQLEEEAAKPPEPEKPVSPPPIESKHRSLVQIIYDENRKKAEAAHRI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LEGLGPQVELPLYNQPSDTRQYHENIKINQAMRKKLILYFKRRNHARKQWKGCRYDQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       444;
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
31.7%; Score 4187; DB 1; Length 2453;
Best Local Similarity 41.0%; Pred. No. 5.8e-126;
Matches 1103; Conservative 348; Mismatches 796; Indels 444;
                                                                                                                                                                                                                                                                                                                                                                                                                                        52208B40382F7E6A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                           isoform Short)
                                                                                                                                                                                                                                                                                   COIL (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                      POLY-ALA.
POLY-ALA.
POLY-SER.
Missing (in isoform
/FIG-VSP 003411.
F-T (IN REF. 2).
A -> P (IN REF. 2).
                                                                                                                                                                                                                                                                                  COILED COIL (POTE)
PRO-RICH.
CORNR BOX OF ID1.
CORNR BOX OF ID2.
                                                                                                                                                                                                                                                                                                                                POLY-GLN.
POLY-ALA.
                                                                                                                                                                                                                                                                                                                                                                                                                                         AA; 270640
                                                                                                                                                                                                                                                          482
669
550
616
2077
2281
                                                                                                                                                                                                                                                                                                                                           602
1047
1718
1979
2371
                                                                                                                                                                                                                                                                                                                                                                                                                            2090
                                                                                                                                                                                                                                                                                                                                                                                                                                       2453
                                                                                                                                                                                                                                                                                                                                                                   1713
1968
                                                                                                                                                                                                                                                                                                                                                         1044
                                                                                                                                                                                                                                    DOMAIN
DOMAIN
DNA BIND
DNA BIND
                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                           CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            425
                                                                                                                                                                                                                                                                                                                                                                                                                 CONFLICT
                                                                                                                                                                                                                                                                                             DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8 6 6 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           장염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       र्व क्षे
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ઠે
```

794 871 851 1361

```
2413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2020
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HISEVITODYTRHH-POOLSAPL----PAPLYSFPGASCPVLDLRRPPSDLYLPPPD--- 2189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      --HGAPA-RGSPH----SEGGKRSPEPNKTSVLGGGEDGIEPVSPPEGMTEPG-HSRSA 2240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HGVVMSHPV---GIMPGSASTSV----VTSSEARRDEGEPSPHAGVCKPKLINKSNSRK 2347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AKSPAPGLA--SGDRPPSVSSVHSEGDCNRRTPLTNRVWEDRPSSAGSTPFPYNPLIMRL 2471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ESTROYEGP-----LHHYRSQOESPSPQQOPPLPPSSQSEGMGQVPRTHRLITLAD 2071
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2084
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1760 YLPTAPQPFSSR-HSSSPLSPGGPTHLTKPTTTSSSERERDRDRERDREREKSILTST 1818
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1819 TTVEHAP----IWRPGTEQSSGSSGGGGGSSSRPASHAHQHSPISPRTQDA-LQQR 1874
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1743 ERIAAAPADLYLRPGSEQ------PGRPGSHGYVRSPSP-SVRTQETILJQR 1787
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PSVLHNTGMKGIITAVEPSKPTVLRSTSTSSPV----RPAATFPPATHCPLGGTLDGVXP 1930
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1847 APOMDVSKTKESKHBAARLBENIRSRSAÅVSEOQOLEOKNIEVEKRSVOCVCTSSALPSG 1906
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1978 GSEBRPLV-------PPVSGHATIARTPAKN--LAPHHASPDPPAPASASDPHRE 2024
                                                                                                                                                                                                                                                                                                                         LTSTPRE----IAKSPHSTVPEHHPHPISPYEHLLRGVSGVDLYRSHIPLAFDPTSIPRG 1640
                                                                                                                                                                                                                                                                                                                                                                           ATAMAQRADMLRGLSPRESSLALNYAAGPRGIIDLSQVPHLPVLVPPTPGTPATAMDRLA 1759
IGSPGRIFPPVHPLDVMAD-ARALERACYE----ESLKSRPGTASSSGGSIARGAPVIV 1530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1124 VLHPRPGPRVSPENLVDKSRGSRPGKSPERSHI---PSEPYEPISPPGG---PAVHEKQD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2241 VYPLLYRDGEQTEPSRMGSKSPGNTSQPPAFFSKLTESNSAMVKSKKQEINKKLNTHNRN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ----SPPLSANAFNPLNASASLPAAMPITAADGRSDHTLTSPGGG-GKAKVSGRPSSRK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EPEYNISOPGTEIFNMPAITGTGLMTYRSOAVOEHASTNMGLEAIIRKALMGKYDOWEE-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          --VARPERPRADTGHAFLAKPPARSGLEPASSPSK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1907 KAQPHASVVYSEAGKDKGPPPKSRYEEELRIRGKITITAANFIDVIITRQIASDKDARER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    KTOSKPFSIQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDKGLPKHLEELDKSHLEG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ELRPKQPGPVKLGGEAAHLPHLR-----PLPESQPSSSPLLQT--APGVKGHQRVVTLAQ
                                                                                                                                                            1531 PELGKPRQSPLTYEDHGA-----PFAGHLPRGSPVTMREPTPRLQEGSLSSSK-ASQDRK
                                                                                                                                                                                                                       | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1641 IPLDAAAAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAALEN-RQTIINDYITSQQMHHNT
                                                  1931 TLMEPVLLPKEAPR-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2025
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2085
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2021
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2072
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1634
                                                                                                                                                                                                                                                                      1473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1700
                                                                                                     1418
                                                                                                                                                                                                                                                                                                                                                  1585
                             1477
```

g

ò

ద

g

ò

ð

셤

6 6 6 6 6

à

6 6 6 6 6 6 6 6

à

유 상 a

8 8 8

용

9 6 9

∂

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Corepressor complex the transcriptional repression activity of some nuclear receptors by promoting chromatin condensation, thus nuclear receptors by promoting chromatin condensation, thus nuclear receptors by promoting chromatin condensation, thus nuclear receptors of the basal transcription.

Therefore the complex that contains SIMSAP and histone deacetylases corepressor complex that contains SIMSAP and histone deacetylases corepressor complex that contains SIMSAP and histone deacetylases the retinoid acid receptors (RAR) in the absence of ligand.

SINGELUTIAN LOCATION: Nuclear (RAR) in the absence of ligand.

SINGELUTIAN LOCATION: Nuclear (RAR) in the absence of ligand.

Interacting domains region contains repression domains (RDI, RDZ and RD3). The C-terminal region contains the nuclear receptor and RD3. The C-terminal region contains the nuclear receptor domains (IDI and IDI2).

Interacting domains that are divided in two separate interaction domains (IDI and IDI2).

Sequence referred to as the CORNR box. This motif is required and sufficient to permit binding to unliganded TR and RARS. Sequences Elaming the CORNR box determine nuclear hormone receptor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         [3]
SEQUENCE OF 974-2440 FROM N.A.
MEDLINE=99375328; PubMed=10444336;
Nagaya T., Chen K.-S., Fujieda M., Ohmori S., Richer J.K.,
Horwitz K.B., Lupski J.R., Seo H.;
Localization of the human nuclear receptor co-repressor (hN-CoR) gene between the CMTIA and the SMS critical regions of chromosome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TISSUE=Brain;
MEDLINE=99397452; PubMed=10470851;
Kikuno R., Magase T., Ishkawa K.-I., Hirosawa M., Miyajima N.,
Kikuno R., Magase T., Ishkawa K.-I., Hirosawa M., Majasima N.,
Tanaka A., Kotani H., Nomura N., Ohara O.;
"Prediction of the coding sequences of unidentified human genes. XIV
The complete sequences of 100 new cDNA clones from brain which code
for large proteins in vitro.";
DNA Res. 6:197-205(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SIMILARITY: Contains 1 SANT-A domain.
SIMILARITY: Contains 1 Myb-like domain.
SIMILARITY: BELONGS TO THE N-COR NUCLEAR RECEPTOR COREPRESSORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE=98393736; PubMed=9724795; Wang J., Hoshino T., Redner R.L., Kajigaya S., Liu J.M.; Wang J., Hoshino T., Redner R.L., acute myeloid leukemia, represses "ETO, fusion partner in t(8;21) acute myeloid leukemia, represses transcription by interaction with the human N-COR/mSin3/HDAC1
                                                                                                                                                                                                                                                                                                                                  Eukaryoča; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.
                                            2452
2472 QAGVMASPPPPGLPAGSGPL--AGPH---HAWDEEPKPLLCSQYETLSDSE
                       Proc. Natl. Acad. Sci. U.S.A. 95:10860-10865(1998)
                                                                                                                                                                                     075376; Q9UPV5; Q9UQ18;
16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
10-OCT-2003 (Rel. 42, Last annotation update)
Nuclear receptor co-repressor 1 (N-COR1) (N-COR)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE OF 782-2440 FROM N.A.
                                                                                                                                                                                                                                                                                                                           Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A. TISSUE=Fetal brain;
                                                                                                                                                                                                                                                                                                   NCOR1 OR KIAA1047.
                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=9606;
                                                                                                                                                                       HUMAN
                                                        2406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             complex.
                                                                                                                                                                       g
              ò
```

```
RGKSRSPAPPADKEAFAAEAQKLPGDPPCWTSGLPFPVPPREVIKASPHAPDPSAFSYAP 1079
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1080 PGHPLPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISOG-----MSVQLH 1134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IIGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETA 1305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     L-KRSYESVEGNIKQGMSMRESPVSAPLEGLICRALP--RGSPHSDLKERTVLSGSIMQG 1299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1135 VPYSEHAKAP-VGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEAS 1193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NFYFNYKKRONLDEILQQHKLKMEKERNARRKKKKAPAAASEEAAFPPVVEDEEMEASGV 710
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSSGA----PQ-DSDSSATCSADEV 926
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ::| :::| ::| :|||||||| :|| :|| | |||| :|| ||| ||| :|| ||| ||| ||| :|| ||| ||| :|| ||| ||| :|| ||| :|| ||| ||| :|| ||| :|| ||| ||| :|| ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| 
         VYKDRQFMNVWTDHEKEIFKDKFIQHPKNFGLIASYLERKSVPDCVLYYYLTKKNENYKA 489
                                                                                                                                                                          SCNEBEMVERAEALHASGNEVPRGECSGPATVNNSSDTESIPSPHTEAAKDTGQNGPKPP 770
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      871
                                                                                                                                                                                                                                                                   :
TEEPPPPLPPPPEPISTEPVETSRWTEEEMEVAKKGLVEHGRNWAAIAKWGTKSEAQCK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DKEDLLKEKTDDTSGEDNDEKEAVASKGRKTANSQGRRKGRITRSMANEANSEEAITPQQ
                                                                                                                                                                                                                                                                                                                                                            ATLGADGPPPGPPTPPRRTSRAPIEPTPASEATGAPTPPPAP-PSPS-APPPVVPKEEKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EETAAAPPV-----------EEGEEQKPPAAEELAVDTGKAEEPVKSECTEEAE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              780 VETQVNDSISAETAEQMDVDQQEHSAEEGSVCDPPPATK--ADSVDVEVRVPENHASKVE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GDNTKERDLDRA-----SEKVEPRDEDLVVAQQINAQRPEPQSDNDSSATCSADE-
                                                                                                                                                                                                                                        S-----AELASMELNESSRWTEEEMETAKKGLLEHGRNWSAIARMVGSKTVSQCK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DEAEGGDKNRLLSP----RPSLLTPTGDPRANASPQK--PLDLKQLKQRAAAIPPI----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ------QVTKVHEPPREDAAPTKPAPPAPPPQNLQPESDAPQQPGSSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CNIPIGTPVSGYALYQRHIKAMHESALLE...----EQRQRQEQIDLECRSSTSP
                                                       LVRRSY-RRRGKSQQQQQQQQQQQQQQQQQQQPMPRSSQEEKDEKEKEAEKEEEKPEVEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPA-----DVLYKGTITR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1306 APKRTYDMMEGRVGRAIS-----SASIEGLMGRAIPPERHSPH-HLKEQHHIRGSITQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1359 IPRSYVBAQEDYLRREAKLLKREGTPPPPPPSRDLTEAYKTQALGPLKLKPAHEGLVATV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              KEAGRSIHEIPREEL-----RHTPELPLAPRPLKEGSITQGTPLKYDTGASTTGSKKHDV
                                                                                                                                                                                                                                                                                                                                                                                                                                                            --NEEENPEDSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EGVRLP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1020
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1246
                                                                                                                                                541
                                                                                                                                                                                                                                           601
                                                                                                                                                                                                                                                                                      604
                                                                                                                                                                                                                                                                                                                                                                                                                             711
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        771
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          838
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              946
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1419
                                                                                                                                                                                                                                                                                                                                 651
                                                                                                                                                                                                                                                                                                                                                                               664
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          718
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   829
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            872
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1020
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1190
                                                       482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      733
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              977
                                                                                      g
                                                                                                                                      à
                                                                                                                                                                                      요
                                                                                                                                                                                                                                        ઠે
                                                                                                                                                                                                                                                                               ద
                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                    ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ठे
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 장 : 점
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
         loved. Usage by and for commercial
(See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    429
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    --EGLRA-SADAKKDPAFGGKHEAPSSPISGQPCGDDQNASPSKLSKEELIQSMDRVDR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EPRYPPHSLSYPVQIARTHTDVGLLEYQ - - HHSRDYASHLSPGSIIQPQRRRPSLLSEFQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QSRYPPHSVQYTFPNTRHQQEFAVPDYRSSHLEVSQASQLLQQQQQQQLRRRPSLLSEFH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PGNERSQELHLRPESHSYLPELGKSEMEFIESKRPRLELLPD------PLLRPSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LLATGQPAGSEDLTKDRSLTGKLE-PVSPPSPPHTDPELELVPPRLSKEELIQNMDRVDR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BITMVEQQISKLKKKQQQLEEERAAKPPEPEKPVSPPPIESKHRSLVQIIYDENRKKAEAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HRILEGLGPQVELPLYNQPSDTRQYHENIKINGAMRKKLILYFKRRNHARKQWKQKFCQR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             YDQLMEALEKKVERIENNPRRRAKESKVREYYEKQFPEIRKQRELQERMQSRVGQRGSGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SMSAARSEHEVSEIIDGLSEQENLEKOMROLAVIPPMLYDADQORIKFINMNGLMADPMK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VYKDROVMNMWSEQEKETFREKFMQHPKNFGLIASFLERKTVAECVLYYYLTKKNENYKS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 31.4%; Score 4147.5; DB 1; Length 2440; Best Local Similarity 40.7%; Pred. No. 1e-124; Matches 1099; Conservative 341; Mismatches 785; Indels 475;
                                                                                                                                                                                                       MINICALOR MANAGEMENT OF TRANSCRIPTION CO-repressor activity; TAS. GO; GO:0003714; F:transcription from Pol II promoter; TAS. InterPro; IPR001005; Myb DNA-binding.

Fram: PF00249; myb DNA-binding.

Fram: PF00249; myb DNA-binding; 2.

FROSITE: PS50090; MYB 3; 1.

Nuclear protein; Transcription regulation; DNA-binding; Repress COAIED COIL (POTENTIAL).

DOWAIN 254 312 INTERACTION WITH SIN3A/B.

DOMAIN 259 328 COILED COIL (POTENTIAL).

DOMAIN 299 328 COILED COIL (POTENTIAL).

DNA-BIND 620 670 MYB.

DOMAIN 501 557 COILED COIL (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      60A4D7964D00EDAB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POLY-ALA.
POLY-RO.
POLY-RO.
POLY-SER.
L -> V (IN REF. 2).
W -> R (IN REF. 2).
W -> R (IN REF. 2).
Q -> H (IN REF. 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  INTERACTION WITH ETO.
CORNE BOX OF IDI.
CORNE BOX OF ID2.
       not removed
modified and this statement is not removentities requires a license agreement (sor send an email to license@isb-sib.ch)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POLY-GLN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRO-RICH
                                                                                           EMBL; AF044209; AAC33550.1; -. EMBL; AB028970; BAA82999.1; -.
                                                                                                                                             AB019524; BAA75814.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      270263
                                                                                                                                           EMBL; AB019524; BAA75814
TRANSFAC; T04687; -
Genew; HGNC:7672; NCORI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              617
1816
2059
2267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  603
1035
1712
1963
1014
1509
1561
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1952
1014
1508
1561
1567
2440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    370
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
```

8 6 8 6

ઠે

8 6

8 8 8

8 8 8 8

```
QRVVTLAQHISEVITQDYTRHH-----PQQLSAPLPAPLYSFPGA--SCPVLDLRRPPSD 2182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2183 LYLPPPD-----HGAP-ARGSPH-----SEGGKRSPEPNKTSVLGGGEDGIEPVSPPEGM 2231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1944 DKDARERGSQSSDSS-----SSLSSHRYETPSDAIEVISPASPAPPQEKLQTYQPEVV 1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RLAYLPTAPQPFSSR-HSSSPLSPGGPTHLTKPTTTSSSERERDRDRERERBKSIL 1815
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TSTITVEHAPIWRPGTEQSSGSSGSSGGGGSSSRPASHSHAHQHSPISPRIQDA-LQQR 1874
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1731 AASSDL----YLRPGSEQ------PGRPGSHGYVRSPSP-SVRTQETMLQQR 1771
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PSVLHNTGMKGIITAVEPSKPTVLRSTSTSSPV----RPAATFPPATHCPLGGTLDGVYP 1930
                                                                                                                                                                                                                                                                                                                                                                        1639 RGIPLD-AAAAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAALEN-RQTIINDYITSQQMH 1696
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HNTATAMAQRADMLRGLSPRESSLALNYAAGPRGIIDLSQVPHLPVLVPPTPGTPATAMD 1756
                                                                                                                                                                                                               : | | : | | : | | : | | : | | | : | | | : | | : | | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RSLIGSPGRTFPPVHPLDVMAD-ARALERACYE----ESLKSRPGTASSSGGSIARGAP 1527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2232 TEPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQPPAFFSKLTESNSAMVKSKKQEIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              KKLNTHNRNBPEYNISQPGTEI PUMPAITGTGLMTYRSQAVQEHASTNMGLEAIIRKALM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GKYDQWEE-----SPPLSANAFNPLNASASLPAAMPITAADGRSDHTLTSPGGGG--KA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      KVSGRPSSRKAKSPAP--GLASGDRPPSVSSVHSEGDCNRRTPLTNRVWEDRPSSAGSTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2078 DKSHLEGELRPKQPGPVKLGGEAAHLPHLRPLPESQPSSSPLLQTAPG-----VKGH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1998 KANQAENDPTRQYEGP-----LHHYRP---QQESPSPQQQLPPSSQAEGMGQVPRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -TGHAFLAK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1831 APOMDVSKTKESKHEAARLEENLRSRSAAVSEQQQLEQKTLEVEKRSVQCLYTSSAFPSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PPARSGLEPASS----PSKGSEPRPLVPPVSGHATIARTPAKN---LAPHHASPDPPAPAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2018 ASDPHREKTOSKPFSIQELELRSLGYHGSSYSPEGVEPVSFVSSPSLTHDKGLPKHLEEL
                                          RKLTSTPRE----IAKSPHSTVPEHHPHPISPYEHLLRGVSGVDLYRSHIPLAFDPTSIP
                                                                                                                                                            VIVPELGKPROSPLTYEDHGA----PPAGHLPRGSPVTMREPTPRLOEGSLSSSKASQD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1931 TLMEPVLLPK----EAPRVARPERPRAD-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1963
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1697
                                                                                                                                                                                                                                                                                                                                                                                                                           1517
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1757
                                                                                                                                                                                 1528
                                                                                                                                                                                                                                                                  1463
                                                                                                                                                                                                                                                                                                                                              1583
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1625
```

8 8 8 8 8 8 8 8 8

8 8

요 &

6 6 6 6 6

8 8 8 8

8 8

8 8

6 6 6

8 8 8

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to licenseéisb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TISSUE=Skeletal muscle;

XX MEDLINE=99421707; PubMed=10491148;

XI MEDLINE=99421707; PubMed=10491148;

XI Schuler M.J., Buehler S., Pette D.;

XI Schuler M.J., Buehler S., Pette D.;

XI Enr. J. Buchem. 264:982-988(1999).

XI I Succeptor mRNA isoforms in rat skeletal muscle.";

XI Diochem. 264:982-988(1999).

XI I SUBCINION: Mediates the transcriptional repression activity of some nuclear receptors by promoting chromatin condensation, thus nuclear receptors by promoting chromatin condensation, thus nuclear sasociates with HDACT. Forms a large corepressor complex complex associates with the thyroid (TR) and the retinoid acid receptors (RAR) in the absence of ligand (By similarity).

C. -I SUBCELULAR LOCATION: Nuclear (By similarity).

C. -I DOMAIN: The N-terminal region contains repression functions that are divided into three independant repression domains (RDI, RDZ and RD3). The C-terminal region contains the nuclear receptor-and man RD3). The C-terminal region contains the nuclear receptor-and comains (IDI and ID2).

C. -I DOMAIN: The two interaction domains (ID) contain a conserved sequences of sequences of flanking the CORNR box. This motif is required and sequence referred to as the CORNR box. This motif is required so sufficient to permit binding to unliganded TR and RARS. Sequences of sequences of sequences of sequences of sequences of sequences of sequences.

C. SIDECTION OF STATEM SECEPTOR COREPRESSORS
FPYNPLIMRLQAGVMASPPPPGLPAGSG-PLAGPH---HAWDEEPKPLLCSQYETLSDSE
                         "Aberrant interactions of transcriptional repressor proteins with the Huntington's disease gene product, huntingtin.";
                                                                                                                                                                                                                                                                                                           Eukaryota; Měcazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SIMILARITY: BELONGS TO THE N-COR NUCLEAR RECEPTOR COREPRESSORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; AF124821; AAD32566.1; -.
EMBL; AF059311; AAC14567.1; -.
InterPro; IPR001005; Myb DNA binding.
PROSITE; P550090; MYB 3; PARTIAL.
Nuclear protein; Transcription regulation; DNA-binding; Repressor.
                                                                                                                                                                                                                              10-OCT-2003 (Rel. 42, Last annotation update)
Nuclear receptor co-repressor 1 (N-CoR1) (N-CoR)
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Boutell J.M., Thomas P., Neal J.W., Weston V.J., Duce J.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        101.
102.
                                                                                                                                                   A
                                                                                                                                                            NCR1 RAT STANDARD; PKF; 533 AA Q9WUB5; 070463; 16-007-2001 (Rel. 40, Last sequence update) 16-007-2001 (Rel. 40, Last sequence update) 10-007-2003 (Rel. 42, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        OF
OF
                                                                                                                                                   533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CORNR BOX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Huntington's disease gene product,
Hum. Mol. Genet. 8:1647-1655(1999).
                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
TISSUE=Brain;
MEDLINE=99371771; PubMed=10441327;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   [2]
SEQUENCE OF 476-528 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      59
157
361
                                                                                                                                                                                                                                                                                                norvegicus (Rat)
                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=10116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           48
153
357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Jones A.L.;
              2462
                                                                                                                                                                                                                                                                                                  Rattus
                                                                                                                                                                             8
                                               셤
```

```
pathway.
                                                                                                        J/Su(H)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ++
   2075
                                                                                                                                                                                                                                                            2293
                                                                                                                                                                                                                                                                                                                                               2411
                                                                                                                                                                                                                 2234
                                                                                                                                                                                                                                                                                                                                                                                         RKAKSPAPG--LASGDRPPSVSSVHSEGDCNRRTPLTNRVWEDRPSSAGSTPFPYNPLIM 2469
                                                                24;
                                                                                                                                                                                                                                                                                                     2294 LNTHNRNEPEYNISQPGTEIFNMPAITGTGLMTYRSQAVQEHASTNMGLEAIIRKALMGK 2353
                                                                                                                                                                                                                                      250
                                                                                                                                                                                                                                                                                 309
                                                                                                                                                                                                                                                                                                                          368
                                                                                                                                                                                  142 RTHRLITLADHICQIITQDFARM--QVPSQPSTSTFQTSPSALSSTPV---RTKDSSRYS 196
                                                                                                                                                                                                                                                                                                                                                                     425
                                                                                                                                                                                                                                                                                                                                                                                                              483
                                                                                                                                                  92 MVKANQAENESPQQYEGP-----LITHYRSQQGSPSPQQQPPLPPSSQ-AEGMGQVP 141
                                                                                                          91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE OF 1-112 FROM N.A.
Adachi J., Aizawa K., Akimura T., Arakawa T., Carninci P., Fukuda S.,
Hashizume W., Hayashida K., Hirozane T., Hori F., Imotani K.,
Ishii Y., Itoh M., Kagawa I., Kawai J., Rojima Y., Kondo S., Konno H.,
Koya S., Miyazaki A., Murata M., Nakamura M., Nomura K., Numazaki R.,
                                                                                                                                                                       GHQRVVTLAQHISEVITQDYTRHHPQQLSAPLPAPLYSFPGA--SCPVLDLRRPPSDLYL
                                                                                                                                                                                                                                                                                AVHEKODSMLLLSQRGMDPAEQRSDSRSPGSISYLPYFFTKL-ESTSPMVKSKKQEIFRK
                                                                                                                                                                                                                                                                                                                2016 ASASDPHREKTOSKPFSIQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDKGLPKHLE
                                                                                                      ASDKDARERGSQSSDSS-----SSLSSHRYEAPSDAIEVISPASSPAPPQEKPQTYQPE
                                                                                                                              ELDKSHLEGELRPKQPGPVKLGGEAAHLPHLR-----PLPESQPSSSPLLQTAPG---VK
                                                                                                                                                                                                                 PPPD-----HGAPA-RGSPH----SEGGKRSPEPNKTSVLGGGEDGIEPVSPPEGMTEP
                                                                                                                                                                                                                                     197 PESOSOTVLHPRPGPRVSPENLVDKSRGSRPGKSPERSHI---PSEPYEPISPPOG---P
                                                                                                                                                                                                                                                           G-HSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQPPAFFSKLTESNSAMVKSKKQEINKK
                                                                                                                                                                                                                                                                                                                                              Y-DQWEESPPLSANAFNPLNASASLPAAMPITAADGRSDHTLTSPGGG-GKAKVSGRPSS
                                                                                                                                                                                                                                                                                                                                                                                                     Newberry B.P., Latifi T., Towler D.A., "The RRM domain of MINT, a novel msx2 binding protein, recognizes and
                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MSX2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MINT MOUSE STANDARD;
PRT; 3644 AA.
0505764; 0907U9; 099584; 0907W2;
10-OCT-2003 (Rel. 42, Last sequence update)
115-WAR-2004 (Rel. 43, Last annotation update)
MBXZ-interacting protein (SMART/HDACI associated repressor protein)
MINT OR SHARP OR KIAA0929.
                                                                                                                                                                                                                                                                                                                                                                                                                                  RLOAGVMASPPPGLPAGSG-PLAGPH---HAWDEEPKPLLCSQYETLSDSE 2517
                                                                                                                                                                                                                                                                                                                                                                                                                                              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus
NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FUNCTION, SUBCELLULAR AND INTERACTION WITH
                                                               67;
                                            533;
                                            Length
                                                               Indels
R -> W (IN REF. 2).
A -> V (IN REF. 2).
7 7DF60F82287EC2 CRC64;
                                           DB 1;
                                                    ; Pred. No. 1.4e-16; 68; Mismatches 192;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (ISOFORM 1),
DNA-BINDING,
                                          5.4%; Score 714.5; 38.5%; Pred. No. 1.4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    equlates the rat osteocalcin promoter."; ochemistry 38:10678-10690(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE=99379811; PubMed=10451362;
484 R
497 A'
57794 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE OF 49-3644 FROM N.A. LOCATION, TISSUE SPECIFICITY,
                                                   Best Local Similarity 38.5
Matches 205; Conservative
484
497
533 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    musculus (Mouse)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        regulates the rat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TISSUE=Testis,
CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                2186
                                                                                                                                                                                                                                                                                                                                               2354
                                                                                                        38
                                                                                                                            2076
                                                                                                                                                                       2128
                                                                                                                                                                                                                                                           2235
                                                                                                                                                                                                                                                                                                                          310
                                                                                                                                                                                                                                                                                                                                                                                        2412
                                                                                                                                                                                                                                                                                                                                                                                                             426
                                                                                                                                                                                                                                                                                                                                                                                                                                  2470
                                                                                                                                                                                                                                                                                251
                                                                                                                                                                                                                                                                                                                                                                     369
                                                                                                                                                                                                                                                                                                                                                                                                                                                      484
                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Mus
 FFS
                                                                                                       셤
                                                                                                                                              셤
                                                                                                                                                                      ò
                                                                                                                                                                                          셤
                                                                                                                                                                                                                                  임
                                                                                                                                                                                                                                                           ઠ
                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                            윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ઠે
                                                                                                                            ઠે
                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
```

```
MEDLINE-22483652; PubMed=12594956;

MEDLINE-22483652; PubMed=12594956;

MEDLINE-22483652; PubMed=12594956;

AKITCODA K., Tanigaria K., Tanigaria K., Tun T., Furukawa T.,

A Taniguchi Y., Kurcoka H., Hameda Y., Toyokuni S., Honjo T.;

AT Taniguchi Y., Kurcoka H., Hameda Y., Toyokuni S., Honjo T.;

Taniguchi Y., Rurcoka B. cell development by MINT, a suppressor of Notch Pathway. Baganian pathway.";

Immunity 18:301-312 (2003).

-I- FUNCTION: Essential corepressor protein, which probably regulates different key pathways such as the Notch pathway. Magative requiator of the Notch pathway via its interaction with RBPSUH, and therefore suppresses the transactivation activity of Notch signaling. Blocks the differentiation of precursor B cells into marginal zone B cells. Probably represses transcription via the recruitment of large complexes containing histone deacetylase proteins. May bind both to DNA and RNA.

- I- SUBUNIT: Interacts with NCOR2, HDAC1, HDAC2, RBBP4, MBD3 and MTALLI Interacts with he nuclear receptors RAR and PPRAD.

- Interacts with RAR in absence of ligand. Bind to the steroid receptor RNA coactivator SRA (By similarity). Interacts with MSX2.

- Interacts with RBPGUH; this interaction may prevent the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          <u>.</u>
۵
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Dillinger K.,
W., Liptay S.,
                                                                                                                                                                                                                                                                                                             Ξ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Vago
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S., Makajima D., Nagase T., Ohara O., Koga H.,
"Prediction of the coding sequences of mouse homologues of KIAA ger II. The complete nucleotide sequences of 400 mouse KIAA-homologous cDNAs identified by screening of terminal sequences of cDNA clones randomly sampled from size-fractionated libraries.";
                                                                                                                                                                                                                                                                                                         Sakamoto T., Gotou T., Isagawa Y., Mimura H., Kimura K., Kawaichi
"MINT/spen negatively regulates Notch signaling by inhibiting RBP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ripoll C., Vaç
Mattei M.-G.,
                                                                                                                                                                                    SEQUENCE OF 69-3644 FROM N.A. (ISOFORM 2), AND VARIANTS THR-348;
PHE-762; PHE-773 AND LEU-933.
STRAIN=ICR; TISSUE=Brain;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         is a novel component of the Notch/RBP-Jkappa signalling
    ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             interaction between RBPSUH and NOTCH1.
SUBCELLULAR LOCATION: Nuclear. Associates with chromatin.
ALTERNATIVE PRODUCTS:
Event-Alternative splicing; Named isoforms=2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          two murine and human cochlear
Ohno M., Ohsato N., Saito R., Sakazume N., Sano H., Sasaki I Sato K., Shibata K., Shiraki T., Tagami M., Takeda Y., Waki Watahiki A., Muramatsu M., Hayashizaki Y.; Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE=22261914; PubMed=12374742; Oswald F., Kostezka U., Astrahantseff K., Bourteele S., Zechner U., Ludwig L., Wilda M., Hameister H., Knoechel Schmid R. M.;
                                                                                                                                                                                                                                                                                                                                                                                                                         Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ., Lenoir M.,
Depetris D.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Note=No experimental confirmation available;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IsoId=062504-1; Sequence=Displayed;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          s:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE=97237053; PubMed=9119401;
Crozet F., El-Amraoui A., Blanchard S.
Hamel C., Fizames C., Levi-Acobas F.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE=22579291; PubMed=12693553;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  "Cloning of the genes encoding tw
unconventional type I myosins.";
Genomics 40:332-341(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE OF 2598-3644 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE OF 318-578 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Weil D., Pujol R., Petit C.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBO J. 21:5417-5426 (2002)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          [sold=Q62504-2;
                                                                                                                                                                                                                                                                                                                                                                                 activity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TISSUE=Cochlea;
```

```
1214 -VDDÝBERRSLVHEVGKPPQDVTDDSPPSKKRRTDHVDFDICTKRERNÝRSSRQISEDŠER 1272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1273 TSCSPSVÄHGSFHDDDDPRGSPRLVSVKGSPKGDEKGLPYPNAAVRDDPLKCNPYDSGKR 1332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1159 RAGĢEKPHGGNAEBKLGIDIDHTOSYRKOMEOSRRKORMEMĒIAK---AĒKFGSPKKD-- 1213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        301 RYDQLMBALEKKVERIENNPRRRAKESKVREYYEKQFPEIRK------342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ----QREL--QERMQSRVGQRGSGLS------MSAARSEHEVSE 374
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           375 IIDGLSE------QRINLEK----QMRQLAVI-------PPMLY 400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DADQQRIKFINMN---GLMADPMKVYK-----DRQVMNWSEQEKETFREKFWQHPKNFG 452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         --- DNDEKEAVASKGRKTAN------ 59GRRKGRITRSMANEANS---- 592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   129;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        256 -LYNOP-----SDTRQYHENIKINQAMRKKLILYFKRRNHA---RKQWKQKFCQ-- 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              --RSQBLHLRPESHSYLPELGKSEMEFIESKRPR----LELLPDPLLRPSPLLATGQPAG 131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       706 EYRSDPYEQDIREYSYRQRERERERERFESDRDHERRPIERSQSPVHLRRPQSPGVSPAH 765
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEDITKD--RSI-----TGKLEPVSPP-----SPPHTDPELELVPPRLS 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              KEELIQNMDRVDREITMVEQQISKLK----KKQQQLEEEAAKPPEPEKPVSPPPIESKH 223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               .099 VHAĞLGELTHGSVETQEKKALESKPÖPKÖLÖLENQGPEKEEVRKNYCREREEPAEH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----RRGK------DEKEK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      526 EKEAEKEBEK-------PEVEND------KEDLLKEKT--DDTSGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ------VLYYYLTKKNENYKS---LVRRSYR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    :| | : ::: | | | | | | | RERRIVKEKGEKD------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RSLVQIIYDENRKKAEAAH------RILEGLGPQVELP--------
659 1659 I -> V (IN REF. 3).
669 1669 S -> F (IN REF. 3).
815 1015 A -> V (IN REF. 3).
815 1015 A -> V (IN REF. 3).
820 2027 G -> A (IN REF. 3).
821 2322 A -> V (IN REF. 3).
822 2385 P -> Q (IN REF. 3).
832 2385 P -> Q (IN REF. 3).
832 2385 P -> Q (IN REF. 3).
844 2559 B -> N (IN REF. 3).
850 2505 B -> N (IN REF. 3).
851 2519 D -> N (IN REF. 3).
852 2565 B -> N (IN REF. 3).
853 2565 B -> N (IN REF. 3).
854 2554 T -> S (IN REF. 3).
855 2565 B -> N (IN REF. 3).
856 2566 B -> N (IN REF. 3).
857 2569 B -> N (IN REF. 3).
858 2569 B -> N (IN REF. 3).
859 2569 B -> N (IN REF. 3).
859 2569 B -> N (IN REF. 3).
850 2569 B -> N (IN REF. 3).
851 2569 B -> N (IN REF. 3).
852 2569 B -> N (IN REF. 3).
853 2569 B -> N (IN REF. 3).
854 2554 B -> N (IN REF. 3).
855 2569 B -> N (IN REF. 3).
856 2569 B -> N (IN REF. 3).
857 2569 B -> N (IN REF. 3).
858 2569 B -> N (IN REF. 3).
858 2569 B -> N (IN REF. 3).
859 2569 B -> N (IN REF. 3).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   4.4%; Score 576; DB 1; Length 3644;
llarity 19.1%; Pred. No. 2.18-11;
Conservative 347; Mismatches 1114; Indels 1042;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GLLEYQHHSRDYASHLSPGSIIQPQRR-RPSLLSEFQPGNE----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     453 LIASFLERKTVAEC------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Local Similarity
les 590; Conserv
                         1659
1669
1705
1815
1815
2322
2322
2385
2502
2502
2519
2519
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 224 I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            557
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     343
                      CONFLICT
CON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ∂
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           д
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its way non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TISSUE SPECIFICITY: Highly expressed in testis. Expressed at lower level in brain, lung, spleen, liver and kidney. Weakly expressed in cardiac and skeletal muscles and ovary. In spleen, it is marghnessed in follicular B-cells, while it is weakly expressed in DOMAIN: The RID domain mediates the interaction with nuclear
                                                                                                                                                                                                                                      SIMILARITY: Contains which mediates the interaction with NCOR2, is essential for the repressive activity (By similarity). SIMILARITY: Belongs to the Spen family.
SIMILARITY: Contains 1 RID (receptor interacting) domain.
SIMILARITY: Contains 4 RNA recognition motif (RRM) domains.
SIMILARITY: Contains 1 SPC domain.
CAUTION: Ref. 4 sequence differs from that shown due to multiple frameshifts and conflicts that create stop codons.
CAUTION: Ref. 5 sequence differs from that shown due to what seems the both received interaction of the shown due to what seems the both received interaction of the shown due to what seems the both received interaction of the shown due to what seems the both received interaction of the shown due to what seems the both received interaction of the shown due to what seems the both received interaction of the shown due to what seems the both received interaction of the shown due to what seems the both received interaction of the shown due to what seems the both received in the shown due to what seems the both received in the shown due to what seems the shown 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                to be the presence of intronic sequence in the CDNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WITH MSX2.
WITH RBPSUH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Missing (in isoform 2) /FIId=VSP_008564.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REF.
REF.
REF.
REF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  INTERACTION W
INTERACTION W
ARG-RICH.
TYR-RICH.
TYR-RICH.
ALG-RICH.
ALG-RICH.
PRO-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL, BY726481, -; NOT ANNOTATED CDS.
EMBL, AF156529; AADS5931.1; ALT INIT
EMBL; AB055980; BAB32786.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              , Z78160; CAB01562.1; ALT SEQ., AK122402; BAC65684.2; ALT SEQ. MGI:1891706; Mint.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ٨
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR000504; RNA_rec_mot.
Pfam; PF00076; rrm; 3.
SMART; SM00360; RRM; 3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               721
832
2233
2518
3475
640
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  754
1524
1560
1570
1574
1609
                                                                                                                                                                                                                                receptors.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
DOMAIN
DOMAIN
VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VARIANT
VARIANT
VARIANT
VARIANT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MGD;
```

```
2409 PSSRKAKSPAPGLASGDRPP----SVSSVHSEGDCNRRTPLIN-RVWEDRPSSAGSTPF 2462
                            MEDLINE-99246063; PubMed=10231032;
MEDLINE-99246063; PubMed=10231032;
Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosawa M.,
Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
"Prediction of the coding sequences of unidentified human genes. XIII.
The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro.";
DNA Res. 6:63-70(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FUNCTION, AND INTERACTION WITH RBPSUH.
MEDLINE=22261914; PubMed=12374742;
Oswald F., Kostezka U., Astrahantseff K., Bourteele S., Dillinger K.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDILNE-12087127; PubMed=11867749;
Shi Y., Hon M., Evans R.M.;
"The peroxisome proliferator-activated receptor delta, an integrator
of transcriptional repression and nuclear receptor signaling.";
Proc. Natl. Acad. Sci. U.S.A. 99:2613-2618(2002).
                                                                                                                                                                                                                          MINIT HUMAN STANDARD; PRT; 3664 AA.
096758; Q9H9A8; Q9WWH5; Q9WQ01; Q9Y556;
10-0CT-2003 (Rel. 42, Last sequence update)
10-0CT-2003 (Rel. 42, Last sequence update)
MAX2-interacting protein (SMART/HDACI associated repressor protein).
MINIT OR SHARP OR KIAA0929.
Homo sapiens (Human).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, INDUCTION, RNA-BINDING, AND INTERACTION WITH NCOR2; HDAC1; HDAC2; RBBP4; MBD3; RAR AND MTAIL1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE OF 793-1595 FROM N.A., AND VARIANT PRO-1091.
TISSUE-Embryo, and Teratocarcinoma;
TISSUE-Embryo, and Teratocarcinoma;
TISOUR-Embryo, and Teratocarcinoma;
TISOUR-Embryo, and Teratocarcinoma;
TISOUR-Embryo, and Teratocarcinoma;
Nishikawa T., Nagain K., Sugano S., Takahashi-Fujii A., Hara H.,
Tanase T., Namura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
Arita M., Nabamura Y., Nagahari K., Masuho Y., Oshima A.;
"NEDO human cDNA sequencing project.";
Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                            Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Buteleostomi,
Mammalia, Eutheria, Primates, Catarrhini, Hominidae, Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TISSUB-Liver, and Pituitary;
MEDLINE=21231190; PubMed=11331609;
Shi Y., Downes M., Xie W., Kao H.-Y., Ordentlich P., Tsai C.-C.,
Hon M., Evans R.M.;
"Sharp, an inducible cofactor that integrates nuclear receptor
repression and activation.";
Genes Dev. 15:1140-1151(2001).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Rhodes S., Huckle E.;
Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
                                                                                                                             3300 PVASSISLASRIKTSAQVPPEGEPLQSTQSAQP 3332
                                                                                    PY-NPLIMRLQAGVMASPPPGLPAGSGPLAGP 2494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE OF 2002-3664 FROM N.A.
TISSUE=Brain;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE OF 294-3664 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       [6]
INTERACTION WITH PPARD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=9606;
                                                                                          2463
                                                                                                                                                                                                                                           硆
                                                                                          ò
```

```
RA Sechner U., Ludwig L., Wilda M., Hameister H., Knoechel W., Liptab S., Stange I. B. novel component of the Notch/RBP-Jkappa signalling pathway."

RA MAY CHANGE CAMPAINT IS ANGERONS OF SPOC DOWAIN.

WE WILL SECTION OF SECTION OF SECTIONARY OF SECTIONAR
```

PSVGPPSVTVVTLESAPSALEKTTGDKTV ENYKKRONLDEILOOHKLKMEKERNARRKKKAAPAAASEEAAPPVVEDEEMEASGVSGN	1893 ŚLPLSRTRRRNVRSVYATWGD-HENRSPVKEPVEGPRVTRKRLERELĢEAĀAVPTTPR 1949 981 VHEPPREDAAPTKPAPPAPPP	1204
8 8 8 8 8 8 8 8 8 8	8 8 8 8 8 8 8 8 8 8	8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8
DOMAIN 517 589 RNA-BINDING (RRW DOMAIN 588 715 COLIED COIL (POT DOMAIN 1170 1191 COLIED COIL (POT DOMAIN 1170 1191 COLIED COIL (POT DOMAIN 1496 1529 COLIED COIL (POT DOMAIN 1928 1944 COLIED COIL (POT DOMAIN 1928 1944 COLIED COIL (POT DOMAIN 2201 2707 RID. COLIED COIL (POT DOMAIN 2130 2464 SPC. INTERACTION WITH DOMAIN 2209 2870 INTERACTION WITH DOMAIN 240 325 ARG-RICH DOMAIN 240 325 ARG-RICH DOMAIN 2428 2520 PRO-RICH DOMAIN 2428 2520 PRO-RICH DOMAIN 3220 3482 PRO-RICH DOMAIN 3220 3482 PRO-RICH DOMAIN 3220 3482 PRO-RICH ARG-RICH DOMAIN 320 3482 PRO-RICH ARG-RICH DOMAIN 320 3482 PRO-RICH ARG-RICH ARG	CONFLICT 956 956 G -> D (IN REF. 4). SEQUENCE 3664 AA; 402245 MW; 5228658533E5B27B CRC64; Query Match Best Local Similarity 19.6%; Pred. No. 3.4e-11; Best Local Similarity 19.6%; Pred. No. 3.4e-11; Matches 568; Conservative 349; Mismatches 973; Indels 1011; Gaps 71 EFOPGNERSCELHLRPESHSYLPELGKSEMEFIESKRPRIELLP 1142 ERKSGQEKSHSVNTEEKIGIDIDHTQSYRKQMEGSRRKQQMEMEIAKSEKFGSPKKOV 115 DPLLRPSPLLATGQPAGSEDLTKDRSLTGKLEPVSPPSPPHTDPELELVPPRLSKEELIO	266 YHENIKINQAMRKKLILYPKRRNHARKQMAGMAKIKLSVLNSEDELMRWDSQM 1344 266 YHENIKINQAMRKKLILYPKRRNHARKQMKPCQRYDQLMEALEKKVERIENNPRR 322 1345 KQDAGRFDVSFPNSIIKRDSLRKRSVNDLEPGEVPSDEDGEHKKSHSPRASAL 1398 233RAKESKVREYYEKQFPEIRKQRELQERWQSRVGGRGSGLSMSAAR 367 1399 YESSRLSFLARDREDKLRERDERKSSERNKFYSFALDTTTPDTKALLERAKSLSSSR 1458 248 SEHEVSEIIDGLSEQENLEKQRRQLAVIPPMLYDADQQRIKFINMAGLMADPMKVY 423 249 YESSRLSFLARDREDWRSKEKVDSAPRPIPSWYMKKKIR-TDSEGKMDDKKEDH 1515 240 KCDQVMNWMSCEKETFREKFMQHPKNFGLIASFLERKTVAECVLYYYLTKKNENYKSLV 483 241 KCDQVMNWMSCEKETFREKFMQHPKNFGLIASFLERKTVAECVLYYYLTKKNENYKSLV 483 242 KCDQVMNWMSCEKETFREKFMQHPKNFGLIASFLERKTVAECVLYYYLTKKNENYKSLV 483 244 KCDQVMNWMSCEKETFREKFMQHPKNFGLIASFLERKTVAECVLYYYLTKKNENYKSLV 483 245 KCDQVMNWMSCEKETFREKFMQHPKNFGLIASFLERKTVAECKMLYTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT

1740	3374IGQPGQPPSSKMPQVSQEAKGIQIGVEQPRADFAGFANRFFDFAIQ	QQ
REPEAT 98	ITAADGRSDH	λō
DOMAIN 1982 DOMAIN 41	PPCPPSQ- 337	qa
FT DOMAIN 1398 1403 PT DOMAIN 1436 1442	Δ.	ò
DOMAIN 814 DOMAIN 1340	3287TRTY 3314	qq
DOMAIN 684 DOMAIN 699	2270 AFFSKLTESNSAMVKSKKOEINKKLNTHNRNEPEYNISOPGTEIFNMPAITGTGLMTY 2327	ò
DOMAIN 519 DOMAIN 636	3231 KEAAKTPDAKAAPTPTPAPVPVPVPLPAPAPAPHGBARILTVTPSNQLQGLPLTPP 3286	qα
MIM; 142580; Repeat.		ò
FIR; Genew	3175 SEVLVMQSEYRLHPYTVPRDVRIMVHPHVTAVSEQPRAADGVVKVPPASKAPQQPG 3230	qu
EMBL; M3351 EMBL; Z1502	LGGGEDG 222 :	ò
EMBL; EMBL;	3134PQRASTPQPAPAGVPALASQHPPEEEVHYHLPVARATAPVQ 3174	q
or send		δλ
modifie entitie	3095 QGEVRMNTPTLPSITYSIRPEALHSPR-APLQPQQIEVRA 3133	d
	ivi .	ò
This SW between	3035 SFPRASHPSSTASTALSTNATVMLAAGIPVPQFISSIHPEQSVIMPPHSITQTVSLSHLS 3094	ପ୍ର
	PVSPVSSPSLT	}
fami Nat.	2000 - NNIARFRIANSELEGATE FARTHOLD	ð
	VSTPVTQGGTVKVLTQGINTPPVLVHNQLVLIPSIPVINNCLADFVILLALE AV LENE ELAV LENE ELA LEN	අූ
RA Iris F.J.M., Bougueleret I	1985 1999	à
RP [2] RP SEQUENCE OF 1-1860 FROM N.	ANVATHSTLVLTAQTYNASPVISSVK-ADRPSL-EKPEPIHLS 291.	ପ୍ର
	RPERP	ζ
"A gene pair irom large proline-ric	2807 RVNTSEGVVILISYSGQKTEGPQRISAKISQIPPASAMDIEFQQSVSKSQVKPDS 2860	qq
RA MEDLINE=90192810; Fubmed=2 RA Banerji J., Sands J., Stro	LGGTLDGVYPTLMEPVLLPKEA	δλ
SEQUENCE FROM N.A. TISSUE=T-cell;	2751 GGVTATTGTVTWAGAVIAPSTKCKQRASANENSRFHPGSMPVIDDRPADAG-SGAGL 2806	셤
OX NCBI_TAXID=9606; RN [1]	999999	'n
	2698NVLTGPVNVLTTPVNATVGTVNAAPGTVNAAASAVNATASAVTVTAGAVTAAŠ 2750	qq
	1806 RDREREKSILTSTITVEHAPIWRPGTEQSSGSSGSS 1841	ờ
28-FEB-2003 (Rel. Large proline-ric	2671 PVKGSVTTLKSLVSTPAGPVNVLKGPV	අ
DT 01-FEB-1996 (Rel. 33, Crea DT 01-FEB-1996 (Rel. 33, Last	1746 PIPGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTHLTKPTTTSSSERERDRDRERD 1805	δ
BAT2 HUMAN ST. P48634;	2638TLAKPAPQTLTGLVSALTGLVNVSLVP-VNALKG 2670	q
SULT 8 T2 HUMAN	1686 INDYITSQQMHHNTATAMAQRADMLRGLSPRESSLALNYAAGPRGIIDLSQVPHLPVLVP 1745	ò
	2585 KTAPPVTNNSEIQASEVLVAADKEKVAPVIAPKITSVISRMPVSIDLENSQKI 2637	ପ୍ଧ
DD 3473 LPHTEFOPAPKQDSSPH	1626 SHIPLAPDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAALENROTI 1685	ò
Qy 2493 GPHHAWDEEPKPL	: : : :: : : : : : : : : : : : : :	Ωp
Qy 2434 HSE-GDCNRRTPLINKW : : :: : Db 3423 QAETGPTSFPSPVSVSM	1579 ASQDRKLTSTPREIAKSPHSTVPEHHPPLSPYEHLLRGVSGVDLYR 1625	ò
7270		

```
s copyright. It is produced through a collaboration itute of Bioinformatics and the EMBL outstation tics Institute. There are no restrictions on its stitutions as long as its content is in no way ment is not removed. Usage by and for commercial ense agreement (See http://www.isb-sib.ch/announce/ense@isb-sib.ch).
2434 HSE-GDCNRRTPLTNRVWEDRPSSAGSTPFPYNPLIMRLQAGVMASPPPPGLPAGSGPLA 2492
                         i=2156268;
rominger J.L., Spies T.;
man major histocompatibility complex encodes
ins with multiple repeated motifs and a
                                                                                                                                                                                                                                                   eated)
st sequence update)
st amotation update)
in BAT2 (HLA-B-associated transcript 2) (G2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    L., Prieur S., Caterina D., Primas G., riguez-Tome P., Claverie J.-M., Dausset J.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             id a potential new member of the NF kappa B
ase HLA class III segment.";
93).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Limited to cell-lines of leukemic origin.
                                                                                                                                                                                                                                                                                                                                                       data; Craniata; Vertebrata; Buteleostomi;
stes; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POLY-PRO.
GIN-RICH.
POLY-PRO.
POLY-PRO.
POLY-PRO.
POLY-PRO.
POLY-PRO.
4 X 57 AA TYPE A REPEATS.
1-1.
1-2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    main.";
.S.A. 87:2374-2378(1990).
                                                                                                                                                                                                                PRT; 2142 AA.
                                                                                 LLCSQ 2509
                                                                                                                 HLTSQ 3493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1; -.
1; JOINED.
1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N.A.
|=8499947;
               ò
```

653 QQHQQQSAPPIPUPPSEPQPVII-GAVPAPGAP	1742 VLVPPTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTHLTKPTTTSS 1793
4 3 4 3 4 3 4 3 4 3 4 3 4 3 4 3 4 3 4 3	8 8 8 8 8 8 8 8 8 8
PT DOWALN 337 549 2. X TYPE B REPEATS.	

```
2371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2431
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   --GDCNRRTPLTNRVWEDRPSSAGSTPFPYNPLI--M 2469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2030 IGRAELHPVELKPPQDYQKLSSNIGGPGSSRTPPTGRSFSGLNSRLKATPSTYSGVFRTQ 2089
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1842 ITSKQWRLRPGTPSLHPYRSQPLYLPPGPAPPSALLSGVALKGQFLDFSTWQATELGKLP 1901
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PGNTSQPPAFF-----SKLTESNSAMVKSKKQEINKKLNTHNRNEPEYNISQPGTEI 2313
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1902 AGGVLYPPPSFLYSPAFCPSPLPDTSLLQVRQ-------DIPSP-SDF 1941
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1942 YSTPLQPGGQSGFLPSGAPAQQ-----MLLPWVDSQLPVVNFGSLPPAPP---PAPPPL 1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1727 DRGTERGPIRPS-HRPGPPVQF-GTSDKDSDLRLVVGDSLKAEKELTASVTEAIPVSRDW 1784
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ---KTSVLGGEDGI- 2222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1785 ELLP---SAAASAEPOSKNLDSGHCVPEPSSSGQRLYPEVFYGSAGPSSSQISGGSHGLS 1841
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SSDPHFEEPGPMVRGVGGTPRDSAGVSPFPPKRRERPPRKPELLQEESLPPPHSSGFLGS 1572
                                                                                                                                            2372 NASASLPAAMPITAADGRSDHTLITSPGGGGKAKVSGRPSSRKAKSPAPGLASGDRPPSVS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2314 FNMPAITG--TGLMTYRSQAVQEHASTNMGLEAIIRKALMGKYDQWEESPPLSANAFNPL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SLLPVGPALQPPSLA------RPPPAPATRVLPSPA------RPPPAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Treisman J.E., Luk A., Rubin G.M., Heberlein U.; "eyelid antagonizes wingless signaling during Drosophila development and has homology to the Bright family of DNA-binding proteins."; (2)
                                                                                                                                                                                                                           2038 LRSLGYHGSSYSPEGVEPVSPVSSPSLTHDKGL-----PKH----LEELDKSHLEG
                                                                                                                                                                                                                                                                                                                                                 2085 ELRPKOPGPVKLGGEAAHLPHLRPLPESOPSSSPLLOTAPGVKGHORVVTLAQHISEVIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QDYTRHHPQQLSAPLPAPLYSFPGASCPVLDLRR--------PPS-DL
                                                                                                         --ELE
                                                                                                                                                                                                                                                                                        1633 QSDSGVDLSGDSQVSSGPCSQRSSP----DGGLKGAAEGPPKRPGGSSPLNAVPCEGPPG
                                                                                                                                                                                                                                                                                                                                                                                                             ------PGPIGTER----SOR----T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  OSA DROWE STANDARD; PRT; 2716 AA.

081N94; O61603; Q9VEG7;
10-0CT-2003 (Rel. 42, Last sequence update)
10-0CT-2003 (Rel. 42, Last sequence update)
10-0CT-2003 (Rel. 42, Last annotation update)
10-0CT-2003 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, TISSUE SPECTFICITY, AND DEVELOPMENTAL STAGE.
MEDLINE=97415319; PubMed=9271118;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RLOAGVMASP-----PPPGLPAGSGPLAGPHHAWDEBP 2502
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2128
                                                                                                               -PDPPAPPASASD-----PH----REKTQSKPFSIQ-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2090 RVDLYQQASPPDALRWIPKPWERTGPPPREGPSRR-AEEP
SSPSKGSEPRPLVPPVSG----HATIARTPAK-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2184 YLPPPDHGAPARGSPHS---EGGKRSPEPN-
                                                                                                                                                                                                                                                                                                                                                                                                                       1689 ЅЕРР̀ЯКР̀РР̀АРНОĠDRКЕLР̀КЕОР̀LР̀-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2432 SVHSE-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=7227;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2223
                                                               1513
                                                                                                                        2009
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              OSA_DROME
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          음
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                           ò
                                                                                                                        ò
```

```
MEDLINE_20195006; PubMed=10'3113;

RAD Adams M.D., Celniker S.E., Holt R.A., Brans C.A., Gocayne J.D.,
RA Adams M.D., Celniker S.E., Holt R.A., Hoskins R.A., Galle R.F.,
RA Adams M.D., Celniker S.E., Richards S., Ashburner M., Henderson S.N.,
Sutton G.G., Worthan J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Baxter E.G., Helt G., Champe M., Feliffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley B.M.,
RA Ballew R.M., Basu A., Barden B.P., Bhandari D., Bolshakov S.,
RA Ballew R.M., Bouck J., Bayraktaroglu L., Beasley B.M.,
RA Besson K.Y., Bence P.V., Berman B.P., Chalter P., Chandra I.,
RA Burtis K.C., Busan D.A., Buller H., Cadieu E., Center A., Chandra I.,
RA Beblos B., Delchar A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Beblos B., Delcher A., Deng C., Perriac C., Perriac R., Parischmann W.,
RA Burtis K.C., Busen D.A., Helman T.J., Hernandez J.R., Houck J.,
RA Bablos B., Kanney B.A., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Helman T.J., Wei M.-H. Ibegwam C.,
RA Harris N.L., Harvey D.A., Helman T.J., Wei M.-H. Ibegwam C.,
RA Hostin D., Houston K.A., Houland T.J., Wei M.-H. Ibegwam C.,
RA Harklov G., Milshan N.V., Mobarry C., Morris J., Murny D., Marklov G., Milshan N.V., Mobarry C., Morris J., Murny D.M., Nelson D.R.,
RA Merkulov G., Milshan N.V., Mobarry C., Morris J., Murny D.M., Nelson D.R.,
RA Reinert K., Remington K.B., Winskern D.R., Percele T., Shen H.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Spier R., Spradling A.C., Stapleton M., Strong R., Wang S., Yao Q., Zhen B.,
RA Mang S.M., Woodage T., Worley K.C., Wu D., Yang S., Zhu X., Smith H.O.,
RA Schong R.M., Worley K.C., Wu D., Yang S., Zhu X., Smith H.,
RA Schong R.M., Worley R.M., Rangenden D.R., Morley R., Shong 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           K., Clampa P.J., Brook A., Dyson N.;
for modifiers of E2F in Drosophila melanogaster.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Kal A.J., Mahmoudi T., Zak N.B., Verrijzer C.P.;
"The Drosophila brahma complex is an essential coactivator for the trithorax group protein zeste.";
Genes Dev. 14:1058-1071(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               an ARID-domain protein that
chromatin-remodeling factor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AND IDENTIFICATION IN A BRAHMA COMPLEX WITH BRM AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DNA-BINDING, AND IDENTIFICATION IN A BRAHMA COMPLEX WITH BRM AND MEDLINE-20069333; PubMed=10601025; Collins R.T., Furukawa T., Tanese N., Treisman J.E.; Grassociates with the Brahma chromatin remodeling complex and promotes the activation of some target genes."; EMBO J. 18:7029-7040(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                IDENTIFICATION IN A BRAHMA COMPLEX WITH BRM; OSA; M
BAP55; BAP60 AND BAP47, AND FUNCTION AS COACTIVATOR
MEDLINE-20270023; PubMed-10809665;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FUNCTION, AND DEVELOPMENTAL STAGE.
MEDLINE=99112962; PubMed=9895321;
Vazquez M., Moore L., Kennison J.A.;
The trithorax group gene osa encodes
genetically interacts with the brahma
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE=99403006; PubMed=10471712; Staehling-Hampton K., Ciampa P.J., "A genetic screen for modifiers of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FUNCTION AS A COREPRESSOR.
MEDLINE=20573925; PubMed=11124806;
                                   MEDLINE=20196006; PubMed=10731132;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Development 126:733-742(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Genetics 153:275-287(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Science 287:2185-2195(2000)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       regulate transcription."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2
```

```
811111111118
                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                          ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                       MEDLINE=22515897; PubMed=12629041;
MEDLINE=22515897; PubMed=12629041;
MEDLINE=22515897; PubMed=12629041;
MICHALOR EV., Vanolate L., Biryukova I., Ramain P.;
Mentaler-promoter communication mediated by Chip during
Pannier-driven proneural patterning is regulated by Osa.";
Genes Dev. 17:591-596(2003).

-I-FUNCTION: Trithorax group (trxG) protein required for embryonic
segmentation, development of the notum and wing margin, and
photoreceptor differentiation. Required for the activation of
genes such as Antp, Ubx and Eve. Binds to DNA without specific
genes such as Antp, Ubx and Eve. Binds to DNA without specific
genes activate to promoters by
promoter-specific proteins. Essential component of the Brahma
complex, a multiprotein complex which is the equivalent of the
yeast SWI/SNF complex and acts by remodelling the chromatin by
catalyzing an ATP-dependent alteration in the structure of
nucleosomal DNA. This complex can both serve as a transcriptional
coactivator or corepressor, depending on the context. Acts as an
essential coactivator for Zeste, which recruits the whole complex
to specific genes. In contrast, it acts as a corepressor for Wg
target genes, possibly via an interaction with Pan and Gro. It
also acts as a negative required by Pan and Chi. Also represses E2f
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -!- SUBGNIT: Component of the Brahma complex, which is composed of Brm, Osa, Mor. Sari/Bap45, Bap111/Dalao, Bap55, Bap60 and Bap47. Interacts with Par and Chi via its EHD domain.
-!- SUBCELLULAR LOCATION: Nuclear.
-!- TISSUE SPECIFICITY: Ubjquitcusly expressed in early embryo. In thist instar larvae, it is ubjquitcusly expressed in a band just antenna imaginal disks, with a stronger expression in a band just antenna imaginal disks, with a stronger expression in a band just anterior to the morphogenetic furrow.
-!- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically.
-!- DOMAIN: The ARID domains mediates the binding to DNA.
-!- SIMILARITY: Contains 1 EHD (Eyelid homology) domain.
-!- SIMILARITY: Contains 1 EHD (Eyelid homology) domain.
-!- CAUTION: Ref.2 (AAF5545) sequence differs from that shown due to erroneous gene model prediction.
                "Osa-containing Brahma chromatin remodeling complexes are required for the repression of wingless target genes."; Genes Dev. 14:3140-3152(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CG; GO:0005634; Cinucleus; IDA.
GG; GO:0005634; Cinucleus; IDA.
GG; GO:0005677; F:DNA binding; IDA.
GG; GO:0005670; F:DNA binding; IDA.
GG; GO:0004549; P:regulation of transcription; IDA.
GG; GO:0007379; P:segment specification; IMP.
GG; GO:0007379; P:segment specification; IMP.
GG; GO:0016055; P:Wing margin morphogenesis; IMP.
GG; GO:0016055; P:Wing receptor signaling pathway; IMP.
InterPro; IPR006031; XYPPX.
InterPro; IPR006031; XYPPX.
Pfam; PF01388; ARN.
InterPro; IPR01389; ARN.
InterPro; IPR006031; XYPPX.
Pfam; PF01389; ARN.
Ffam; PF01389; ARN.
Ffam; FP01389; ARN.; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Transcription regulation; DNA-binding; Activator; Repressor; Chromatin regulator; Nuclear protein; Developmental protein. DOMAIN 997 1111 ARID. EHD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; AF053091; AAC06254.1; -.
EMBL; AE003718; AAF55457.1; ALT_SEQ.
EMBL; AE003718; AAN13750.1; -.
Collins R.T., Treisman J.E.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FlyBase; FBgn0003013; osa.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PIR; T13049; T13049.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              activation.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           + +
```

```
1022 KSRSPAPPADKEAFAAEAQKLPGDPPCWTSGLPFPVPPREVIKASP-HAPDPSAF-SYAP 1079
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     980 KVHEPPREDAAPTKPAPPAP---PPPQNL------QPESDAPQQPGSSPRG 1021
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1080 PGHPLPL-------GLHD---TARPVLPRPPTISNPPPLISSAKHPSVLERQIGAI 1125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1126 SQGMSVQLHVPYSEHAK-----APV--GPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQ 1177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1178 AGPPESLGVPTAQEASVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADV 1237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SGPPHETAAPKRTYDMMEGRVGRAISSASIEGLMGRAIPPE--RHSPHHL---KEQHHIR 1352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            806
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         670 KLKMEKERNARRKKKKAPAA-----ASEEAAFPPVVEDEEMEASGVSGNEEEMVEEAEA 723
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   78 GPGGPPGPDPAAVMHYHHLHQQQQQHPPPPHMQQQQHHGGPAPPPPGGAPEHAPGVKEEY 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             --TPPPAAP-----GE 841
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    138 THLPPPHPHPAYGRYHADPNMDPYRYGQPLPGGKPP----QQQQPHPQQQPPQQPGPGGS 193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             342 EQKPPAAEELAVDTGKAEEPVKSECTEEAEEGP-----AKGKDAEAAEATAEGALKAEK 895
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    194 PNRPPQORYIPGOPPQGPTPTLNSLLQSSNPPPPPPQHRYANTYDPQQAAASAAAAAAQQ 253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       254 QOAG-GPPPPGHGPPPPQHQPS-----PYGGQGGWAPPPRPYSPQLGPSQQYR 301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 --PSLLTPTGD---PRAN-----ASPQKPLDLKQLKQRAAAI------PPIQVT 979
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           302 TPPPTNTSRGQSPYPPAHGQNSGSYPSSPQQQQQQQQQQQQQQQGGPVPGGPPPGTG 361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           362 Q--QPPQQNTPPTSQYSPYPQRYPTPPGLPAGGSNHRTAYSTHQYPEPNRPWPGGSSPSP 419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   582 ASNSASSASNSPQQT-----PPPAPPPNQGMNNMATPPPPP---QGAAGGGYPMPP 629
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    675 TGPP----PPPTSQAGAGGANSMPSGAQAGGYPGRGMPNHTGQYPPYQWVPPSPQQTVP 729
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         724 LHASGNEVPRGECSGPATVNNSSDTESIPS-----PHTEAAKDTGQNGPKPPATLG 774
                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   472 PSPHQELIGQNSNDSSSGGAHSGMGSGPPGTPNPQQVMRPTPSPTGSS------GSR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4 KIKSPQTQQQQGGAPAATPPSAGAAPGAATPPT--------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ------BGPPTPNNNSNNGSDPSIQQQQNVAPHPYGA-----PPPPGS--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            775 ADGPPPGP-------PTPPRRTSR----APIEPTPASEATGAP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            420 GSGHPLPPASPH-HVPPLQQQPPPPPHVSAGGP---PP---SSSPGHAPSPSPQPSQAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SMSPAVAQNHPISRPASNOSSSGGPMQQPPVGAGGPPPMPPHPGMP-GGPPQQQSQQQQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1238 LYKGTITRIIGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   630 HMHGGYK--MGGPGQSPGAQGYPPQQPQ------QYPPGNYPPRPQYPPGAYA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 896 KEGGSGRATTAKSSGAPQDSDSSATCSADEVDEAEGGDKN-----RLLSPR---
                                                                                                                                                                                                                                                                                                                                                                               Query Match 3.8%; Score 503; DB 1; Length 2716; Best Local Similarity 20.6%; Pred. No. 3.3e-09; Matches 440; Conservative 195; Mismatches 776; Indels 724;
                                                                                                                                                                                                                                   1169 1169 V -> G (IN REF. 1).
1795 1795 M -> T (IN REF. 1).
2637 2637 G -> E (IN REF. 1).
2716 AA, 284063 MW, BFAE76CB51C7C675 CRC64;
GLN-RICH.
GLY-RICH.
GLY-RICH.
GLN-RICH.
GLN-RICH.
HIS-RICH.
SER-RICH.
MISSING (IN REF. 1).
V -> G (IN REF. 1).
G -> E (IN REF. 1).
                                 873
1453
1751
1745
2624
2716
61
1169
1795
                                                                                           1271
1730
2589
2625
                                                                                                                                                                                                              61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      40
                                                                                                                                                                                                                                      CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             807
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 942
                                                                                                                                                                                                              CONFLICT
                                                                                                                                                                                                                                                                                                   CONFLICT
                                                                                                                                                                                                                                                                                                                           SEQUENCE
                                                                                                                         DOMAIN
                                                                                           DOMAIN
                                                                                                                                                                             DOMAIN
                                    DOMAIN
                                                               DOMAIN
```

		δ	2343 EAI
ò		q	1527
qq		ò	2403
ò	LKEGSITQGTPLKYDTGASTTGSKKHD	Q	1569 PPP
Dβ	HKGGYGGSPTPPQG-PQGYGNGPTGM	δλ	2430 VSS
ò	ACYEESLKSRPGTASSSG	qq	1629
οp	HPGMPMGPPHHMGPPHGPTNMGPPTSTPPQSQMLQGGQPQGGASG	δ	2477A
ර සි	1520 GSIARGAPVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSPVTWRRPTPRLQEGSLSSS 1577	qq	1682 PYA
ò	1578 KASQDRKLTSTPREIAKSPHSTVPEHHPHPISPYEHLLRGVSGVDLYRSHIPLAFDPTSI 1637	RESULT MAPA RA	JLT 10 A RAT
Ωp		a S	MAPA RAT P34926;
δλ		TO	01-FEB-1994 01-FEB-1994
QQ	NWPRPAGSPQVFNH	TO O	16-OCT-2001 Microtubule-
δλ		GN E	MAP1A. Rattus norve
තු	LKAFMEEKKIFIIA	886	
à i	PPTPGTPATAMDRIAYLPTAPQPFSSRHSSSPLSFPGGFTHLIKFTI	S O S	NCBI_TaxID=1
å í	CFILSANVEDDIRE IIIVAAA	RP	SEQUENCE FRC TISSUE=Brair
ර් ස්	CHEDROD DPLATIOOVEAGSKKKTAKAASVPSPG	R RX	MEDLINE=9235 Langkopf A.,
g ;	HSHSTORESEDEN	RT	"Microtubule one messenge
g 43		<u>-</u>	J. Biol. Che -!- FUNCTION bridging
λ̈́o	1892 PSKPTVLRSTSISPVRPAATFPPATHCPLGGTLDGVYPTLMEPVLLPKEAPRVARPERP 1951	388	-!- SUBUNIT
qa .	1175 PDYATAGQMQRPPSQNNPQTPHPGAAAAVAAGDNISVSNPFEDP-IAAGGG 1225	មួម	-!- TISSUE !
ò		ខម	
qq	OAAGOHOO	88	
δλ	SIQELELRSLGYHGSSYSPEGVEPVS	ខ្លួន	
qu		 	
ò	HLPHLRPLPESQPSSS		This SWISS-
Ωp		388	between th
λ	PLLQTAPGVKGHQRVVTLAQHISEVITV :	388	
셤	P-YPSQPGAYGQYGSSDQYNAIGPPGQPFGQCFG	888	entities re
<u>გ</u>		288	EMBL; M8319
8 8		¥ % E	Microtubule
. q		FT	DOMAIN
ò	占	FT	REPEAT REPEAT
QQ	1491SQPYPGYNARPQIYGAWQSGTQQYRPQYPSSPAPQN 1526	FT	REPEAT

```
S-PROT entry is copyright. It is produced through a collaboration the Swiss Institute of Bioinformatics and the EMBL outstation an Bioinformatics Institute. There are no restrictions on its non-profit institutions as long as its content is in no way and this statement is not removed. Usage by and for commercial requires a license agreement (See http://www.isb-sib.ch/announce/remail to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Seenger RNA.";
1. Chem. 267:1656(1992).
2. Chem. 267:1656(1992).
3. Che
                                                                                                                                                                   2429
                                                                                                                                                                                                                   -------wggapp--rgaapppgapHgppiQQP--AGVaQwDQHRYPPQQGPP 1568
                                                                                                                                                                                                                                                                                                                               ----- 2476
                                                                                                                                                                                                                                                                                                                                                                                  2343 BAIIRKALMGKYDQWEESPPLSANAFNPLNASASLPAAMPITAADGRSDHTLTSPGGGGK 2402
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----AKVSGRPSSRKAKSPA-----PGLASGD------RPPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          egicus (Rat).
Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Sutheria, Rodentia, Sciurognathi, Muridae, Murinae, Rattus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hammarback J.A., Mueller R., Vallee R.B., Garner C.C.; le-associated proteins 1A and LC2. Two proteins encoded in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MAP1 LIGHT CHAIN LC2.
LYS-RICH (BASIC).
11 X 3 AA REPEATS OF K-K-[DE].
1.
3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AMPPPPSQGVGQQVGQGP---PGGMMSQKPPPM 1713
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ASPPPP----GLPAGSGPLAGPHHAWDEEPKPL 2505
                                                                                                                                                                                                                                                                                                                                            SVHSEGDCNRRTPLTNRVWEDRPSSAGSTPFP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT; 2774 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        197, A43359.

Le; Repeat; Phosphorylation.

2774 MAPI LIG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 n;
155629; PubMed=1379599;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      96; AAB48069.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          496
541
338
417
422
426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 OM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                10116;
                         8
```

ATTAKSSGA TSTEEATEP LKQLKQRAA LKGQQPHKD - SRSPAPPA 	1077 YAPIPCHIF 1108 LISSAKHPSVLER 1108 LISSAKHPSVLER 1130 LKSETRQQKGQILDER 1145VGPVTMGLPLP 1200 LGSVPCGSITKGIPS 1429 LQCT 1201 LGSVPCGSITKGIPS 1429 LQCTCRS 1251SPSRLDRCRS 1251RESPLDR 1265 SVTQCSKEDGRS 1333 RAIPEPERHSP 1593 ESSPAEGSKAREOEK 1394 PPPPPPSRDLT 1431 EELRHTPELPL 1431 EELRHTPELPL 1433 EBSPEGEVRYWRDRDIT 1431 EELRHTPELPL 1703 EQSITPLQHTPRSPW 1485 -PPVHPLDWADDARA 1485 -PPVHPLDWADDARA 168	1718SSL
6 8 6 8 6 8 6	品	ò
T REPEAT 427 429 5. T REPEAT 431 433 6. T REPEAT 436 438 7. T REPEAT 440 442 8. T REPEAT 440 444 9. T REPEAT 449 451 10. T REPEAT 539 541 11. D SEQUENCE 2774 AA; 299526 MW; 3DEF74427BA9D7D7 CRC64; Query Match 3.7%; Score 492.5; DB 1; Length 2774; Best Local Similarity 20.5%; Pred. No. 7.2e-09; Matches 553; Conservative 278; Mismatches 1002; Indels 861; Gaps 132;	7.7 BRSORLHURPBSHSTLPBLGKSEMBTIESTRRPRLELLDPDLLRPSPLLATGOPAGSEDJT 136 3.9 DIGGRGUILK-BKTSIGLE 3.99 DIGGRGUILK-BKTSIGLE 4.1 KDAKUGRIKL-BKTSIGLE 4.2 ROGEKELSSERPTPP-ACKALRPDLKPTFEVRKTLYKAKACHGRIKTURGAL193 4.3 KORGEKELSSERPTPP-ACKAACHARGELTAK	
	8 4 8 4 8 4 8 4 8 4 8 8 8 8 8 8 8 8 8 8	

ATTAKSSGAPQDSDSSATCSADEVDEAEGGDGNRILLSPRPSLITPTGDPRANASPOKED 1STERITE POWDEVLRFTOGSLSPRINGSLSPUSYVSPTT

```
2252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2647
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2516
                                                                                                                                                                                                                                        1791 ISSSERERDRDRERDRDREREKSILTSTTTVEHAPIWRPGTEQSSGSSGSSGGGGGSSSR 1850
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1911 ATFPPATHCPLGGTLDGVYPTLME------PVLLPK----EAPR-----V 1945
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2291 PAPAPAPGLP-GDLGDGTLPCRPECTGELTKKPSPFLSPSGDHEANGPGETSLNPPGFVT 2349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1995 ARTP---AKNLAPH--HASP---DPPAPPASASDPHREKTOSKPFSIQELEL-----RS 2040
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EAGPOGCATDPRPHCGELSPSFLNPPLPPS-----TDDSDLSTEEARLAGKGGRRR 2456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LGYHGSSYSP----EGVEPVSPVSSPSLTHDKGLPKHLEELDKSHLEGELRPKQPG---P 2093
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             V-KLGG-----EAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQRVVTLAQHISEVITQ 2145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2146 DYTRHHPQQLSAPLPAPLYSFPGASCPVLDLRRPPSDLYLPPPDHGAPARGSPHSEG--- 2202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2551 DVCMADPEGLSS-----ESGRVERLREKGR-----PGRRAPGRAKPASPARRL 2593
                                                                                                                                                PASHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGIITAVEPSKPTVLRSTSTSSPVRPA 1910
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ARPERPRADIGHAF------LAKPPARSGLEPASSPSKGSBPRPLVPPVSGHATI 1994
2108 WPETEAYSSLSSDSHLGSVRPSLDFPASAFGFSSLQPAP--PQL--PSPAEPRSAPCGSL 2163
                                                                                                      ---DR-LAYLPTAPQPFSSRHSS------SPLSPGGP--THLTKPTT 1790
                                                                                                                                                                                                                                                                                                           2222 PALSE------GSSSEATTPVISSVAERFP---PGLE----AAEQSAEGLGSGKE 2263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         "Structure and expression pattern of human ALR, a novel gene with strong homology to ALL-1 involved in acute leukemia and to Drosophila trithorax.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2594 DIRGKRSPTPGKGPVDRTSRTVPRPRSTPSQVTSAEEKDGHSPMSKGLV-----NGLKA
                                                                                                                                                                                                                                                                                                                                                                                                                                             2264 SAAHS-LWDLTPLSPAPSASLDLAPA-----PA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2517 VDKAGGVSGTHHPRPGHDPPPTPLPDPRPSPP------RP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2203 ---GKRSPEPNKTSVLGGGEDGIEPVSPPEGMT----EPGH---SRSAVYPLLYRDGEQT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2648 GSTALGSK---GGSGPPVYVDLAYIPNHCSGKTADQDFFRRVRASYYVVSGNDP 2698
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2253 EPSRMGSKSPGNTSQPPAFFSKLTESNSAMVKSKKQEINKKLNTH----NRNEP 2302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            O14686; O14687;
10-OCT-2003 (Rel. 42, Last sequence update)
10-OCT-2004 (Rel. 43, Last sequence update)
15-WAR-2004 (Rel. 43, Last annotation update)
Myeloid/lymphoid or mixed-lineage leukemia protein 2 (ALL1-related
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
MEDLINE=97388474; PubMed=9247308;
Prasad R., Zhadanov A.B., Sedkov Y., Bullrich F., Druck T.,
Rallapalli R., Yano T., Alder H., Croce C.M., Huebner K., Mazo A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Homo sapiens (Human).
Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT; 5262 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Oncogene 15:549-560(1997)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     protein).
MLL2 OR ALR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HUMAN
                                                                                                                                                                                                                                                                                                                                                                                        1851
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1946
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2041
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2094
                                                                                                                1756
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ID ACCOUNT TO A SELECT THE ACC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                  ద
                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                          원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation—the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                       TISSUE-Cervical carcinoma;
MEDINE=22371496; PubMed=12482968;
GOO Y.-H., Sohn Y.C., Kim D.-H., Kim S.-W., Kang M.-J., Jung D.-J.,
Kawak E., Barlev N.A., Berger S.L., Chow V.T., Roeder R.G.,
Azorga D.O., Meltzer P.S., Suh P.-G., Song E.J., Lee K.-J., Lee Y.C.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                               ISOId=014686-3; Sequence=VSP 008560;
-!- TISSUE SPECIFICITY: Expressed in most adult tissues, including a variety of hematcipoidetic cells, with the exception of the liver.-!- MISCELLANEOUS: This gene mapped to a chromosomal region involved in duplications and translocations associated with cancer.-: SIMILARITY: Contains 5 PHD-type zinc fingers.-: SIMILARITY: Contains 1 post-SET domain.-: SIMILARITY: Contains 1 RING-type zinc finger.-: SIMILARITY: Contains 1 RING-type zinc finger.-: SIMILARITY: Contains 1 RING-type zinc finger.-: SIMILARITY: Contains 1 SET domain.
                                                                                                                                "Activating signal cointegrator 2 belongs to a novel steady-state "Activating signal cointegrator 2 belongs to a novel steady-state complex that contains a subset of trithorax group proteins.";
Mol. Cell. Biol. 23:140-149[2003).
-!-FUNCTION: May be involved in transcriptional regulation.
-!- SUBMIT: Belongs to the ASC-2/NCOA6 complex (ASCOM), which contains ASC-2/NCOA6, the retinoblastoma-binding protein RBQ-3/RBBPS, alpha- and beta-tubulins, the trithorax group proteins MLL2 and MLL3, and ASH2/ASCL2 (Probable).
-!- ALTERNATIVE PRODUCTS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TAS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MIM; 602113; -.

RIM; 602113; -.

RO; GO:0005634; C:nucleus; TAS.

RO; GO:00037048; P:cranscription factor activity; TAS.

RO; GO:0007048; P:oncogenesis; TAS.

RO; GO:0006366; P:transcription from Pol II promoter; TAS.

RINEEPPO; IPR003889; FYrich N.

RINEEPPO; IPR003889; FYrich N.

RINEEPPO; IPR006118; Recombinase.

RINEEPPO; IPR006118; Recombinase.

RINEEPPO; IPR001214; SET.

RINEEPPO; IPR00154; SET.

RINEEPPO; IPR00155; Znf_PHD.

RINEEPPO; IPR001841; Znf_Ting.
                                                                                                                                                                                                                                                                                                                                                                                                                             IsoId=014686-2; Sequence=VSP_008563, VSP_008559;
                                                                                                                                                                                                                                                                                                                                                 Event=Alternative splicing; Named isoforms=3;
INTERACTION WITH ASC-2/NCOA6 CONTAINING COMPLEX
                                                                                                                                                                                                                                                                                                                                                                        Name=1;
IsoId=014686-1; Sequence=Displayed;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; AF010403; AAC51734.1; -.
EMBL; AF010404; AAC51735.1; -.
ENR; T03454; T03454.
PIR; T03455; T03455.
Genew; HGNC:7133; MLL2.
GN; GO:000534; C:nucleus; TAS.
GO; GO:000534; C:transcription f
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PROSITE; PS50868; POST_SET; 1. PROSITE; PS50280; SET; 1. PROSITE; PS61359; ZE PHD_1; 5. PROSITE; PS50016; ZE_PHD_2; 5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam; PF00628; PHD; 5.
Pfam; PF00886; SET; 1.
SMART; SM00541; FYRN; 1.
SMART; SM00141; FYRN; 1.
SMART; SM00149; HMG; 1.
SMART; SM00149; PHD; 7.
SMART; SM00184; RING; 3.
SMART; SM00184; STNG; 3.
                                                                                                                                                                                                                                                                                                                                                                                                                Name=2;
                                                                                                                            Lee J.W.;
```

226 276 PHD-TYPE 1. 229 274 RING-TYPE 1. 273 273 274 PHD-TYPE 2. 274 275 PHD-TYPE 3. 1102 1155 PHD-TYPE 4. 1102 1259 244 PHD-TYPE 4. 2246 245 PHD-TYPE 4. 2397 2436 COLLED COLL (POTENTI) 2397 2436 COLLED COLL (POTENTI) 2397 2436 COLLED COLL (POTENTI) 2398 2809 COLLED COLL (POTENTI) 341 3701 COLLED COLL (POTENTI) 342 46 46 1. A REPEATS OF 4. 446 47 3. COLLED COLL (POTENTI) 456 466 1. A REPEATS OF 4. 555 559 7. COLLED COLL (POTENTI) 568 8. COLLED COLL (POTENTI) 469 470 3. ARG-RICH. 571 525 56 8. 572 559 7. 573 570 9. COLLED COLL (POTENTI) 574 56 8. 575 570 570 6. 576 570 6. 577 571 572 573 570 6. 578 579 9. COLLED COLL (POTENTI) 579 570 570 6. 570 670 670 670 670 670 670 670 670 670 6	<i>ò</i>	
1102 1155 11529 1284 1229 1284 1229 1284 1284 2805 2397 2436 22974 3001 3437 3476 3437 3476 3437 3476 3437 3476 3437 3476 3437 3476 3437 3476 3437 3476 3437 3476 3437 3476 346 469 468 663 663 663 664 663 667 663 667 663 667 673 688 1015 1105 1105 1105 1106 1106 1107 1107 1108 1109 1109 1109 1109 1109 1109 1109	1	VSPP
11102 11155 11102 11155 11229 11289 12346 5242 2397 2436 2974 3071 3437 3477 3476 3437 3476 3437 3476 3437 3476 3437 3476 348 2889 2986 4689 468 4689 663 667 663 667 663 667 663 667 663 667 663 667 663 667 663 667 673 558 674 668 1015 1105 1105 1105 1105 1105 1105 1105	90	1286 VKPVAPVAPP-
1229 1284 5242 5246 5246 2397 2397 2436 2294 2397 2436 2436 2437 442 442 443 446 460 468 469 473 469 469 469 469 469 469 469 469		195 KKOOO
2342 2347 2347 2347 2346 2347 2348 2347 3437 3449 346 346 346 346 346 346 346 346	-	
2397 2436 29788 2889 29788 2889 3286 3342 3437 3476 3437 3476 445 469 460 468 460 468 460 673 555 559 564 568 663 667 663 667 663 667 663 667 663 667 673 589 674 649 674 649 675 649 675 649 676 673 677 689 677 689 678 679 679 679 679 679 679 679 679 679 679 679 679 679 679 679 679 679 670 679 670 679 670 679 670 679 671 679	අ _ර	1335 KRRQRRGRLGLI
29788 29788 3342 3342 3342 3342 3342 3342 3342 33	δ	248 LGPQVELPLYNC
33286 33342 3428 3342 3621 4287 4265 4287 456 466 460 464 460 464 460 464 461 464 462 668 552 552 552 552 564 568 618 667 618 622 623 667 624 631 645 649 645 649 647 641 645 649 647 641 648 640 306 4085 306 4085 306 4085 306 4085 4085 4085 4085 4085 4086 4086 4086 4087 4087 4088 4088 4089 4089 4089 4089 4089 4089	ପ୍ର	: : 1395 VSPDVE-PGKEI
3621 4265 442 446 460 460 460 464 460 464 464		
4265 4287 439 466 460 464 460 464 460 464 460 464 460 508 504 508 504 508 564 568 603 613 618 622 622 632 623 643 623 667 623 667 645 649 645 649 647 631 11015 11053 11015 11	3	300 QK
4439 4439 4460 4400 4400 504 4400 504 4400 504 508 508 508 508 508 508 508 608 608 608 608 608 608 608 608 608 6	qa	1442 QAEVLSGDGQPI
460 469 473 469 473 469 473 469 504 504 508 609 618 609 618 622 627 627 623 623 645 645 645 645 645 647 645 649 647 648 540 529 649 647 649 649 641 648 540 580 580 580 580 580 580 580 580 580 58	-	337 PPRIRKOPRIO
469 469 469 504 504 508 508 509 509 618 620 618 621 622 623 623 624 625 631 645 649 649 649 649 649 649 649 649	<i>(</i> 3	
5.55 5.56 5.57 5.57 5.59 5.69 6.13	q 0	1497 LQEAFFGKELLE
521 555 555 568 573 573 573 573 573 573 663 663 663 663 663 673 663 1015 1015 1015 1015 1022 1015 1023 1032 2587 2703 2587 2703 2587 2703 2587 2703 2587 4000 3966 4085 4090 4090 4090 4090 4090 4090 4090 409	ò	392 LAVIPPMLYDAL
555 559 573 573 573 574 573 618 627 627 627 627 627 629 627 629 629 1012 1025 1025 1032 1032 1032 1032 1032 1032 1032 1032 1032 1032 1032 1032 1032 1033	7	
573 . 577 582 . 586 609 . 613 618 . 622 627 . 631 645 . 649 663 . 667 229 . 326 374 . 922 1015 . 1023 1025 . 1035 1032 . 2547 2586 . 4000 3966 . 4000 3966 . 4000 396 . 4000 4634 . 4702 4634 . 4702 567 . 5703 568 . 5703	q _Q	1547 KGDDGPDIADEE
582 582 583 618 627 627 645 645 649 663 1015 1015 1015 1025 1015 1025 1015 1025	8	446 QHPKNFGLIASE
618 627 627 645 645 663 1012 1012 1012 1012 1012 1012 1012 102 10	i	
627 645 645 645 645 645 645 649 649 1015 1015 1015 1015 1015 1015 1015 101	Q	1588 PGTPGEGMLSS
645 649 663 326 374 922 1105 1053 11832 2351 2587 2767 2587 277 2587 277 2587 4000 3966 4000 3966 4000 3966 672 41454 1454 4949 4949	ò	506 QQQQQPMPRSSC
229 229 374 374 374 374 322 335 2536 2537 2537 2537 2537 2547 2587 2587 2587 2797 306 4085 4085 4085 4085 4085 4085 4085 4085	i	
374 922 11015 11055 11015 11055 11832 2351 2587 2703 2587 2703 3966 4005 4634 4702 4634 4702 306 672 306 672 4949 4949	gg —	1608LPK
1015 1053 11015 1053 11832 2351 2587 2703 2986 4000 4685 4702 4634 4702 306 672 1 305 672 1 454 1454	8	548EKTDDTSC
1832 2351 2536 2547 2536 4000 3966 4000 3966 4000 306 4000 306 672 1 306 672 1 454 1454 4949 4949	<u>'</u>	
2536 2547 2587 2703 2986 4000 3966 4000 3966 4000 305 4702 1 305 672 1 1454 1454	qq —	1660 LPSSSPMDSYPC
2587 2703 2986 4000 3966 4000 3966 4000 1305 305 1000 1000 1000 1000 1000 1000 1000 1	ò	598 POOSAELASMEI
2986 4000 3966 4085 4634 4702 10 306 672 10 1454 1454 1454 4949	7	
1396 4085 4634 4702 10 306 672 1454 1454 1454 4949	q _Q	1712 PTTEGEGDGLSY
306 672 306 672 1454 1454 4949 4949		
306 672 1454 1454 4949 4949	λο 	643 SKTVSQCKNFYF
306 672 1454 1454 4949 4949	qa	1765SRCKQIMK
1454 1454 4949		
1454 1454 4949	λο —	703 EEMEASGVSGNE
1454 1454 4949	qq	1811 KQTKVGDIARKI
SPLSEPERSTRICTORPERSPERSPRICE PLANSINGER PERSPERSPRICE PROPERSPERSPRICE PROPERSPERSPRICE PROPERSPERSPRICE PROPERSPERSPRICE PROPERSPERSPRICE PROPERSPRICE PROPERS PROPERSPRICE PROPERS	d	
PPEASRLFPPRESPLSPPEESPLSPPEASRLSPPPED PPEASRLSPPEESPLSPPEESPLSPPEESPEESPEESPEESPEESP	ò	
	qu	1863 SADGFLKPPA-
1454 1454 B 4949 R R		
1454 1454 4949	₹ 	797 TPASEATC
1454 1454 4949 4949	ପ୍	1918 PPTYPPYPSPT
1454 1454 4949 4949	ė	
1454 1454 4949	-	O TO THE PARTY OF
1454 1454 4949 4949	qa	1973RCPSLDNLAV
1454 1454 4949	<u></u>	Admiranopolitika Apa
/FTId=VSP_008559. 1454 1454 B -> BGBT (in isoform PTId=VSP_008560. 4949 4949 R -> H (in dbSNP:3782 PTId=VAR_017115.	<i>x</i>	
1454 1454 K BGBT (1n 18010rn 4949 4949 K -> H (in dbSNB:3782 /FTId=VAR (1n dbSNB:3782 /FTId=VAR (207115.	di -	2015 EELGASSP
4949 4949 R -> H (in dbSNP:3782 / FICH AND	\$ \frac{1}{2}	
/FTIG=VAR 017115.	Š	. 946 IFIGUERANASE
	qa	2064 EPQ-SPGLGLRP
SEQUENCE 5262 AA; 5641/1 MW; 268/C/4CAD41/E44 CRC64;	è	
Query Match 3.6%; Score 470.5; DB 1; Length 5262;	₹ —	- ND-3-3-3-3-3-3-3-3-3-3-3-3-3-3-3-3-3-3-3

```
PGEAGLEGSEPSDALGPDDKKDGDLDTDELLKGEGGVEHMECEIKLEGP 1394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       : : | | : : | | ... | | ... | | ... | | | ... | | ... | | | ... | | | ... | | ... | | ... | | ... | | ... | | ... | | ... | | ... | | ... | | ... | | ... | | ... | | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ..
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               : : : : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 20.3%; Pred. No. 6.5e-08; ative 284; Mismatches 891; Indels 1019; Gaps 144;
                                                                                                                                                                                                                                                                                                                                                                                                                           GLCQSPFLDSRERGGFFSPEPGEPDSPWTGSGGTTPSTPT-----T 1711
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VPESPGVGGKASEPLLSP-----PPFGESRK-----ALEVKK 2014
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PSYGPPNLGFVDSPSSGTHLGGLELKTPDVFKAP-----LTPRASQV 2063
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ||:|
RPQEP------PPAQALAPSPFSHPDIFRPGSYTDPYAQPPLT 2109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              JEEEAA-KPPEPEKPVSPPPIESKHRSLVQIIYDENRKKAEAAHRILEG- 247
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        IQPSDTRQYHENIKINQAMRKKL-----ILYFKRRNHARKQWKQKFC 299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ADQQRIK-----FINMNGLMADPMKVYKDRQVMNMWSEQEKETFREKFM 445
                                                                                                                                                                                                                                                                                  ITDPELELVPPRLSKEE-----LIQNMDRVDREITMVEQQISKLK 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PERMOSR ----- VGORGSGLSMSAARSEHEVSEIIDGLSEQENLEKQMRQ 391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FLERKTVAECVLYYYLTKKNENYKSLVRRSYRRRGKSQQQQQQQQQQQQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QEEKDEKEKEKE --- AEKEEE --- KPEVENDKEDLLK ----- 547
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3G-----EDNDEKEAVASKGRKTANSQGRRKGRITRSMANEANSEEAIT 597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FNYKKRONLDEILQQHKLKMEKERNARRKKKKAPAAASEEAAFPPVVED 702
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BEEMVEEAEALHASGNEVPRGECSGPATVNNSSDTESIPSPHTEAAKDT 762
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LGADGPPGPPT----- PPRRTSRA------PIEP---- 796
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VDT-----GKAEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEK 895
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               KSSGAPQ----DSDSSAT-----CSADEVDEAEGGDKNRLLSPRPSLL 945
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ILNESS--RWTEEE----METAK------KGLLEHGRNWSAIARMVG 642
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GAPTPPPAPPSPSAP------PPVVPKEEKEEETAAAPPVEEGE 841
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POKPLDLKOLKORAAAIPPIQVTKVHEPPRED------AAPTKP 994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SDLDRISTEE
```

OY 1956 GHAFLAKPPARSGLEPASSPSKGSEPRPL'		Oy 2338 TUMGLEALIRKALMGKYDQWEESPPLSAN Db 3329	SULT 12 SN_DROME SPEN_DROME OBSX83; C9NHN: 10-CCT-2003 (f) 10-CCT-2003 (f) 10-CCT-2003 (g) Split ends program or CG184; SPEN OR CG184; Drosophila melbukaryota; Mel Bukaryota; Mel Bukaryota; End Bukaryota; End Bukaryota; End Rephydroidea; I	RP USAGE, TISSUE SPECIFICITY, DEVELOPMENT RP D57. RX MEDLINE=20025936; PubMed=10556062; RA Wiellette El., Harding K.W., Mace K.P. RA McGinnis W.; Harding K.W., Mace K.P. RT 'spen encodes an RNP motif protein that ro repress the development of head-lik RT trunk."; Pevelopment 126:5373-5385(1999). RM [2] RN P SCUENCE FROM N.A. (ISOFORM 4). TISSUE=Embryo; RX MEDLINE=20157049; PubMed=10655223; RA Rebay I., Chen F., Hsiao F., Kolodziej
2110 PRPQPPPESCCALPPRS-LPSDPFSRVPVSPQSQSSQSPLTPRPLSAEAFC 2161 1044 GDPPCWTSGLPFPVPPREVIKASPHAPDPSAFSYAPPGHPLPLG-LHDTARPVLPRP 1099 2162PSPVTPRFQSPDPYSRPPSRPQSRDFAPLHKPPRPQP 2199 1100 PTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHV-PYSE-HAKAPVGPVT 1149	STRVPSDSAITYRGSITHGTPADVLY	2473 LVGLPPSKLSGPILGPGSFPSDDRLSRPPPPAT 2505 1388 PPSRDLTEAYKTQALGPLKLKPAHEGLVATVKEAGRSIHEI 1428 2506 PSSMDVNSRQLVGGSQAFYQRAPYPGSLPLQQQQQLWQQQQATAATSMRFAMSARFPST 2565 1429 PREEL-RHTPELPLAPRPLKEGSITQGTPLKYDTGASTTGSKKHDVRSLIGSPG 1481 2566 PGPELGRQALGSPLAGISTRLPGPGEPVPGPAGPAGFIELRHNVQKGLGPG 2617 1482 RTFPPVHPLDVMADARALERACYEESLKSRPGTASSSGGSIARGPVIVPELGKPRQSPL 1541 2618 TPPP	1542 TYEDHGAPFAGHLPRGSPVTMREPTPRLQEGSLSSSRASODRKLTSTPREIAKSPHSTVP 1601 2635 SEDFHRLAPEGLRGLAVSGLPPQKPSAPPAP-ELNNSLHPT 2674 1602 EHHPHPISPYEHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLDAAAAYYL 1651 2675PHTKGPTLPTGLELVNRPPSSTEIGRFN-PLALEAGKLPCEDPELDDDFDAHKALED 2730 1652 PRHLAPNPTYPHL-YPPYLIRGYPDTAALERROT	1740 LPVLVPPTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTHLTKPTTTSSSERE 1797 1
8 8 8 8 8 8	3	8 8 8 8 8 8	3 8 3 8 3 8 3 8	8 8 8 8 8 8

APL-PAPLYSFPGAS------CPVLDLRR 2178 RGSPHSEGGKRSPEPNKTSVLGGGEDGIEPVSP 2227 :| : :: : 3289 HRKSRKALCAKQRTAKKAGREFPEADAEKL-- 3289 ANAFNPLNASASLPAAMPITAADGRSDHTLTSP 2397 VLALSP-SQSPRLLTKLPGQLLPG---HGLQPP 3369 DRPPSVSSVHSEGDCNRRTPLTNRVWEDRPSSA 2457 - 3379 RDGEQTEPSRMGSKSPGNTSOPPAFFSKLTES 2278 ISQPGTEIFNMPAITGTGLMT-YRSQAVQEHAS 2337 LVPPVSGHATIARTPAKNLAPHHASPDPPAPP 2015 HGSSYSPEGVEPVSPVSSPSLTHDKGL-PKHL 2074 . | : | : | 3073 HLPHLRPLPESOP--SSSPLLQTAPGVKGHQR 2131), FUNCTION, ALTERNATIVE PROMOTER INTAL STAGE, AND MUTANTS E9 AND ej P.A., Kuang B.H., Laverty T., .A., Ronshaugen M.R., Wang F.Y., hat interacts with Hox pathways ike sclerites in the Drosophila poda; Insecta; Pterygota; achycera; Muscomorpha; la. 5560 AA. PL1; Q9VPL2; update) on update)

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Holt R.A., Babuburner M., Henderson S.N., Button G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Batadon R.C., Rogers Y.-H.C., Blazej R.G., Champpe M., Prefeifer B.D., RA Brandon R.C., Rayers T.H.C., Blazej R.G., Champpe M., Prefeifer B.D., RA Amidon R.C., Baxendell W.D., Zhang Q., Chen L.X., Miklos G.L.G., Maril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D., Ballew R.M., Basu A., Barendal J., Bayraktaroglu L., Beasley E.M., Beson K.Y., Bennes P.V., Burner B.D., Bhandari D., Bolshakov S., Burtis R.C., Busam D.A., Bhiller H., Cadieu E., Center A., Chanfra I., Rodeon K., Doug L.E., Downes M., Dayan-Rocha S., Dunkov B.C., Dunn P., Bodson K., Doug L.E., Downes M., Dayan-Rocha S., Dunkov B.C., Dunn P., Bodson K., Doug L.E., Downes M., Dayan-Rocha S., Dunkov B.C., Dunn P., Bodson K., Doug L.E., Downes M., Dayan-Rocha S., Dunkov B.C., Busan D., Harris M.L., Harvey D.A., Heiman T.J., Hernandez J.R., Heischmann W., R. A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M. A., Harris M.L., Harvey D.A., Heiman T.J., Hernandez J.R., Necknin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C., Jalali M., Kalubh F., Karpen G.H., Ke Z., Kennison J.A., Ketchmu K.A., Anders B. Levister B. Kodira C.D., Kraft C., Mcrais J., Moshrefi A., Montko M., Murphy B., Murphy L., Murny D.M., Nelson D.L., Nelson D.L., Ratkerlow G., Mishian N.V., Mobary C., Morris J., Puri V., Reese M.G., Spier E., Spradling A.C., Stapleton M., Stupski M.P., Smith T., Spier E., Spradling A.C., Stapleton M., Stupski M., Stang S., Yang S., Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J., Walle M., Weille M., Weille M., Weille M., Weille M., Weille M., Weille M., Shubski M., Stang K., Shen H., Spier E., Spradling A.C., Stapleton M., Stupski M., Stang S., Yang S., Yang S., Yang S., Yang S., Weille M., We Suh C., Voas M., Williams A., Rubin G.M.;
"A ganetic screen for novel components of the Ras/mitogen-activated
protein kinase signaling pathway that interact with the yan gene of
Drosophila identifies split ends, a new RNA recognition motif-REVISIONS, AND ALTERNATIVE SPLICING.
MEDLINE=22426069; PubMed=12537572;
Mista S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
Hradecky P., Huang Y. Kaminker J.S., Millburn G.H., Prochnik S.E.
Smith C.D., Tupy J.L., Whitfield E.J., Bayraktarcglu L., Berman B.)
Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drygdale R.A.,
Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,
Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M., MEDLINE=20171275; PubMed=10704397; Kuang B.H., Wu S.C.-Y., Shin Y.-A., Luo L., Kolodziej P.A.; Split ends encodes large nuclear proteins that regulate neuronal cell fate and axon extension in the Drosophila embryo."; Development 127:1517-1529(2000). SEQUENCE FROM N.A. (ISOFORM 1), SUBCELLULAR LOCATION, AND TISSUE Annotation of the Drosophila melanogaster euchromatic genome: Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002). SEQUENCE FROM N.A., AND TISSUE SPECIFICITY. MEDLINE=20196006; PubMed=10731132; Genetics 154:695-712(2000). containing protein."; systematic review."; STRAIN=Berkeley; rissum=Embryo; Lewis S.E.; SPECIFICITY

STRAIN=Berkeley; TISSUE=Embryo; SEQUENCE OF 424-2002 FROM N.A.

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -[- FUNCTION: Probable corepressor protein, which regulates different key pathways such as the EGF receptor and Mg pathways. Involved in neuronal cell fate, survival and axon guidance, cell cycle regulation and repression of head identity in the embryonic trunk. May act with the Hox gene Deformed and the EGF receptor signaling pathway. Positive regulator of the Mg pathway in larval tissues but not in embryonic tissues. May act as a transcriptional corepressor protein, which repress transcription via the recruitment of large complexes containing histone deacetylase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Name=4; Synonyms=SpenS; Isolate and Part and Par
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -!- SUBCELLULAR LOCATION: Nuclear.
-!- ALTERNATIVE PRODUCTS:
Event=Alternative promoter;
Comment=2 isoforms, 1 (shown here) and 2, are produced by use of
                     Stapleton M., Carlson J.W., Broketein P., Yu C., Champe M., George R.A., Guarin H., Kronmiller B., Pacleb J.M., Park S., Wan K.H., Rubin G.M., Celniker S.E.;
"A Drosophila full-length cesource.";
Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -i- SIMILARITY: Belongs to the Spen family.
-i- SIMILARITY: Contains 3 RNA recognition motif (RRM) domains.
-i- SIMILARITY: Contains 1 SPOC domain.
-i- CAUTION: It is uncertain whether Met-1 or Met-7 is the initiator.
                                                                                                                                                                                                                                                                                                                                                          "A screen for modifiers of cyclin E function in Drosophila melanogaster identifies Cdk2 mutations, revealing the insignificance of putative phosphorylation sites in Cdk2."; Genetics 155:233-244(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Chen F., Rebay I.; "split ends, a new component of the Drosophila EGF receptor pathway, regulates development of midline glial cells."; Curr. Biol. 10:943-946(2000).
                                                                                                                                                                                                                                                                  MEDLINE=20253107; PubMed=10790398;
Lane M.E., Elend M., Heidmann D., Herr A., Marzodko S., Herzig A.,
Lehner C.F.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FUNCTION ON WG PATHWAY.

PUNCTION ON WG PATHWAY.

MEDLINE-25668976, PubMed=12783785;

Lin H.V., Doroquez D.B., Cho S., Chen F., Rebay I., Cadigan K.M.;

"Splits ends is a tissue/promoter specific regulator of Wingless
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Name=3; Synonyms=SpenL;
IsoId=Q8SX83-3; Sequence=VSP_008567;
Note=Produced by alternative splicing of isoform 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       the Swiss Institute of Bioinformatics and the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Name=2;
IsoId=Q8SX83-2; Sequence=VSP 008565, VSP 0085
Note=No experimental confirmation available;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Event=Alternative splicing; Named isoforms=4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 IsoId=Q8SX83-1; Sequence=Displayed;
MEDLINE=22426066; PubMed=12537569;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE=20414403; PubMed=10959845;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FUNCTION ON EGF RECEPTOR PATHWAY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Development 130:3125-3135(2003)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  alternative promoters;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        zygotically.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Name=1
                                                                                                                                                                                                                                      FUNCTION
```

```
3830 QQHMQQFMHQQMIQRQQHMQQQQLHGQSQQITSAPQHQMHQQHQAQQQQ
              508
                                                               ò
                                                                                                               8
                                                                                                                                       a
                                                                                                                                                                   8
                                                                                                                                                                                      임
                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                              쉱
                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ठे
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2421 -DIPDDCPNVSMYDK--VKARSCKNMQRQAEE-----KKIKAKFSQLKQSRAK---- 2465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    :| ||: :| || | 202
2743 KERKEKKREKLRNMTEATVPNSPTTNDTSSEKLSKEERHRLKKSKKSKSMDNSCNTKIYN 2802
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     tch 3.5%; Score 466.5; DB 1; Length 5560; al Similarity 18.4%; Pred. No. 9.1e-08; 600; Conservative 373; Mismatches 1159; Indels 1125; Gaps 142;
                                                                                                                                                                                                                                                                                                                                                                                                                                                     ------KKRS---TSYDGDSDIEFEDRQHRNSGSSSFHGRYPGLSSSDDDDDEET 2511
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MNG------LMADPMKVYKDRQ------VMNWWSEQEKETFREKFMQHPKNFG 452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            256 ----INQAMRKKLI--LY 283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              308 ALEKKVERIEN-----NPRRRAKESKVREYYEKQFPEIRKQ-RELQERMQSRVGQR 357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QENLEKQMRQL----AVIP--PMLYDADQQRI-------KFIN 411
                                                                                                                                                                                                                                                                                                                                                                                                                                       91 YLPELGKSEMEFIESKRPRLELLPDPLLRPSPLLATGQPAGSEDLTKDRSLTGKLEPVSP 150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----ARKOW---KOKFCORYDOLME 307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GSGLSMSAARSE-----SE 381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LIASFLERK-TVAECVLYYY----LTKKNENYKSLVRRSYRRGKSQQQQQQQQQQQQ
                                                                                                                                                                                                                                                                                                                                                                                      90
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            31 ARTHIDVGLLEYQHHSRDYASHLSPGSIIQPQRRRPSLLSEFQPGNERSQELHLRPESHS
                                                                                                                                                                                                                                                                                                                                                                                                               SH----SDECH----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        151 PSPPHTDPELELVPPRLSKEELIQNMDRVDREITMVEQQISKLKKKQQQLEEEAAKPPEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         211 EKPVSPPPIESKHRSLVQIIYD-----ENRKKAEAAHRILEGLGPQVELP----
                                                                                                                                                                                                                                          Transcription regulation; Repressor; Developmental protein;
Nuclear protein; Repeat; RNA-binding; Coiled coil;
Alternative promoter usage; Alternative splicing.
DOMAIN 554 632 RNA-BINDING (RRM) 1.
DOMAIN 656 730 RNA-BINDING (RRM) 2.
DOMAIN 734 806 RNA-BINDING (RRM) 3.
                              EMBL; AF184612; AAR26299.1; ---
R EMBL; AF221715; AAR34661.1; ALT_INIT.
REMBL; AE003590; AAF51534.2; ---
REMBL; AE003590; AAF51534.2; ---
REMBL; AE003590; AAR51535.2; ---
REMBL; AV094788; AAN10511.1; ---
REMBL; AV094788; AAN1141.1; ---
REMBL; AV094788; AAN1141.1; ---
REMBL; AV094788; AAN1141.1; ---
REMBL; AV094788; AAN1141.1; ---
REMBL; AV09478; AAN1141.1; ALT_SEQ.
REMBL; AV094788; AAN1141.1; ALT_SEQ.
GO; GO:00007411; P:axon guidance; IMP.
GO; GO:0000847; P:axon guidance; IMP.
INTERPRO; IPRO00564; RNA_rec_mot.
REMB.; SWART; SW00360; RRW; 3.
send an email to license@isb-sib.ch)
                                                                                                                                                                                         SMART; SM00360; RRM; 3.
PROSITE; PS50102; RRM; 3.
PROSITE; PS00030; RRM RNP_1; FALSE_NEG.
PROSITE; PS50917; SPOC; 1.
                                                                                                                                                                                                                                                                                                                                                                                                            AATNTSSGKLHHOHHRRSVERKSSRG-
                         EMBL; AF188205; AAF13218.1; -. EMBL; AF184612; AAF26299.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            284 FKRRNH-----
                                                                                                                                                                                                                                                                                                                               Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                              2358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                2391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2572
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2466
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   412
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
8
                                                                                                                                                                                                                                                                                                                                                                                                          엄
                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

	6	SABOUNDS OF THE SABOUNDS OF TH
	508	NDREDLEREKIDDISGEDNDERE 56
	י ע	
	1500	י אַנ
	620	METAKKGLLEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDEILQQHKLKMEKER 677
•	3017	
	678	NARRKKKKAPAAASBEAAFPPVVEDBEMEASGVSGNEBEMVEBABALH 725
•	3043	KKRSKEEKQEKLLQQQRRESLPNVASTSSA-PPTPGKLTVNVQAASKHADLQLDAKH 3098
	726	TES
•	3099	ISSPPVCKPSPSLPCLIGDDDDDALHTPKAKPTTPSSRGNDGLTPS-REK 3147
	786	PRRISRAPIEPTPASEATGAPTPPPAPPSPSAPPPVVPKEEKEEETAAAPPVEEG 8
	3148	PRLISPIPKTPTIANSSTLSTQSAETPVSSGTVISSSALATTPTSSTAAGVSAAPG 3203
	84	BELAVDTGKAEEPVKSECTEEAEGPAKGKDAEAARATAEGALKAEKKEGGS 90
	3204	LDNSPTSASAQCKKKESFIPGFDGQLDDRISESAVQSISAEFNS 3247
	901	
_	3248	TSLLDNIADEPKIPVASPPRATKPLDKLE
	930	RLLSPRPSLLTPTGD
•	3306	TTDYSLDGMDEMSSVNELETPTLVIAEPDEEAALAAKAIETÄGEPASILEE 3356
	3357	PPREDAAPTKPAPPAPPPPQNLQPB1013
	۰ .	OPGSSPRGKSRSPADADKEAFAAEAOKLPGD 104
	3414	: ; PTPRETANIDIPNVESQ 347
	1046	PPCWTSGLPFPVPPRI
•	3473	PKLSNESTPQPSVITKLPFLDTPKTVPAGLPPSPVKIEPPTISKLQQPLQPVQTVLPAP 3532
	1100	1114
•	3533	HSTGSGISANSVINLDLSNVISSCSNTSAASATASASASISFGSPTASQNAMPQASTPKQ 3592
	1115	PSVLERQIGAISQGMSVQLHVPYSEHAKAPVGPVTMGLPLPMD 1157
•	3593	GPITPQQAIRTQSLIMQPPTISIPEQTPHPAVPQWVLSPQSHHPQQP-GTYMVGIRAPSP 3651
	1158	PKKL-APPSGVKQEQLSPRGQAGPPESLGVPTAQEASVLRGTALGSVPGGSITKGI 1212
•	3652	HSPLHSPGRGVAQSRLVGQLSPVGRPWVSQPSPQQQVQQTQQQHALITSPQSSNISPL 3709
	1213	PSTRVITHGTPADV 1237
_	3710	ASPTTRVLSSSNSPTTSKVNSYQPROQVPQQPSPKSVAEVQTTPQLMTIPLQKMTPIQV 3769
	1238	LYKGTITRIIGEDSPSRLDRGREDSLPKGHVIYECKKGHVLSYEG 1282
_	3770	VTVQPQQATQSQVASSPPLGSLPPHKNVHLN
	1283	GMSVTQCSKEDGRSSSGPPHETAAPKRTYDMMEGRVGRAISSA 1325
•	3830	QQHWQQFWHQQMIQRQQHMQQQQLHGQSQQITSAPQHQMHQQHQAQQQQ3878

Db 4859 PQPLKAHVLNREKNIQQQLTPTKQAV Qy 2205 RSPEPRKTSVLGGGEDGIEPVSPEGMTEPGH	
3126 SIEGLMGRAIPPERHSPHILKEQHHIRGSITQQIPRSTV	1806 RDREREKSILTSTTTVEHAPIWRPGTEQSSGSSGSGCGGGSSSRPASHSHAH 1858 1 1 1 1 1 1 1 1 1
8 4 8 4 8 4 8 4 8 4 8 4 8 4 8 8 8 8 8 8	5

| : | : | : | | : : | | : : | | : : | | : : | | | : : | | | : : | | | : : | | | : : | | | : : | | | : : | | | : : | | | : : | | | : : | | | : : | | | : : | | | : : | | | : : | | | : : | | | | : : | | : : | | : : | | : : | | : : | | : : | | : | | : : | | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : | | : : | | : | | : : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | : | | : | | : | : | | : | : | | : | : | | : | : | | : | : | | : | : | | : | : | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | -VMASPPPPGL-PAGSGP---LAGPHHAWD 2499 AVAQPPQH-----APHSGHMLLTDTAGNQ 4906 KSKKQEINKKLNTHNRNEPEYNISQPGTEI 2313 STNMGLEAIIRKA------LMGK 2353 SSVHSEGD--CNRRT------PLIN 2447 :|:| | | | |||| aslhdlgayvsgrrtqspppahqqaspitp 5193 SHSR-----SAVYPLLYRDGEQTE 2253 OLOMPPHGI IMPTHPGMLLQQKLPAHLQPQ 5075 SASLPAAMPITAADGRSDHTLTSPGGGGKA 2403 it presynaptic side of synaptic Vertebrata; Euteleostomi; thi; Muridae; Murinae; Rattus. s-Langhoff C., Voss B., delfinger E.D., Garner C.C.; ein structurally related to protein involved in the f T.C.;
one protein piccolo:
eurotransmitter release."; finger E.D., Garner C.C.; k/DDBJ databases. IS OF ASP-4668; ASP-4674; GLN-4692; ASN-4693 AND RIMS2 and profilin (By ate) pdate) c cytomatrix protein). SRACTION WITH RABACL. 5 AA.

```
H
                                                                                                                                                                                                                                                                                                                                                                                                                                    535
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          589
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              635
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    649
                                                                                                                                                  908
                                                                                                                                                                                                                            907
                                                                                                                                                                                                                                                                                  350
                                                                                                                                                                                                                                                                                                     1004
                                                                                          75
                                                                                                              791
                                                                                                                                                                     184
                                                                                                                                                                                        863
                                                                                                                                                                                                                                                                                                                                                            451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                원
                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                  d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                      쉱
                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                            В
                                                                                                                                                                                                                                                                 ò
                                                                         ₹
                                                                                                             ò
                                                                                                                               셤
                                                                                                                                                 ð
                                                                                                                                                               g
                                                                                                                                                                                      ò
                                                                                                                                                                                                    셤
                                                                                                             This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its worken by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ON->AA: MODERATE INCREASE IN AFFINITY FOR CALCIUM.
                                                                                                                                                                                                                                                                                                                                                                                                 12 X 10 AA TANDEM APPROXIMATE REPEATS OF P-A-K-P-D-D-Q-D-X.
P-A-K-P-C {COTENTIAL}.
C4-TYPE (POTENTIAL).
C4-TYPE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AFFINITY FOR CALCIUM. INCREASE IN AFFINITY FOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CALCIUM.
VM->SS: 10-FOLD INCREASE IN AFFINITY FOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                D-5A: COMPLETE LOSS OF CALCIUM-BINDING
AND CALCIUM-DEPENDENT PHOSPHOLIPID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A: COMPLETE LOSS OF CALCIUM-BINDING CALCIUM-DEPENDENT PHOSPHOLIPID
                                                                                                                                                                                                                                                                      Isoid=09JKS6-2; Sequence=VSP 003930, VSP_003931; DOMAIN: C2 domain 1 is involved in binding calcium and phospholipids. Calcium binds with low affinity but with high specificity and induces a large conformational change. SIMILARITY: Contains 2 C2 domains. SIMILARITY: Contains 1 PDZ/DHR domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A->S: NO EFFECT ON CALCIUM-BINDING
                                                                                                                                                                                                                   GO:0045202; C:synaptic junction; IDA.
GO:0005509; F:calcium ion binding; IDA.
GO:000554; F:calcium-dependent phospholipid binding; IDA.
GO:0005522; F:profilin binding; ISS.
GO:0007010; P:cyroskeleton organization and biogenesis; ISS.
GO:0016080; P:synaptic vesicle targeting; NAS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BINDING ACTIVITY.
V->S: SMALL INCREASE IN AFFINITY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACTIVITY.
MW; SAIBB543201A7450 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                         C2 DOWAIN 1.
C2 DOMAIN 2.
TKPIN -> SKRRK (in isoform
/FTId=VSP_003930.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     P 003930.
(in isoform 2).
ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BINDING ACTIVITY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         003931
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    VV->SS: 10-FOLD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           M->S: INCREASED
                             IsoId=09JKS6-1; Sequence=Displayed;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FTIG=VSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Missing
                                                                                                                                                                                          EMBL; AF138789; AAF07822.2; -.
EMBL; AF227534; AAF63196.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5085 AA; 552702
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              4689
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4689
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4693
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4694
                                                                                                                                                                                                                                                                                                                                                                                                                        547
1033
2362
2362
4536
4752
5059
4880
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5085
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   4668
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              4674
                                                                                                                                                                                                             P04410; 1A25.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             4689
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          4688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          4692
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4694
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 4881
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4668
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4674
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /ARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                       ZN_FING
ZN_FING
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MUTAGEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MUTAGEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MUTAGEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AUTAGEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MUTAGEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MITAGEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MUTAGEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               fUTAGEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
    +
```

```
1560 V------TMREPTPRLQE-----GSLSSSKASQDRKLTSTPREI----- 1592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1333 RAIPPERHSPHHLKEQHHIRGSITQGIPRSYVEAQEDYLRREAKLLKREGTPPPPPPRD 1392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LTBAYKTQALGPL-KLKPAHEGLVATVKEAGRSIHEIPREELRHT----PELPLAPRPLK 1447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EGSITQGTPLKYDTGASTTGSKKHDVRSLIGSPGRTFPPVHPLDVMADARALERACYBES 1507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1059
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PSTRVPSDSALTYRGSITHGTPADVLYKGTITRIIGEDSPSRLDRGREDSLPKGHVIYEG 1272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1115 PSVLERQIGAISQGMSVQ-LHVPYSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLS 1173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1174 PRGQAGPPESLGVPTAQEASVLRGTALGSVP-----------GGSITKGI 1212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              KKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETAAPKRTYDMMEGRVGRAISSASIEGLMG 1332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           KPLDLKQLKQR-----AAAIP-----PIQVTKVHEPPREDAAPTKPAP----PAPPPPQ- 1003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1060 REVIKASPHAPDPSAFSYAPPGHPLPLGLHDTARPVLPRPPTISNP----PPLIS-SAKH 1114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -----TPAKPQPQ-----PP-----TATKPQPQPdPATKPHQQPGLAKPSAQQ 489
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      906
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                KSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRP---SLLTPTGDPRANASP----Q 958
                                                                                                                                                                                                                                     -----PSMHRKQELDSSQ-APQQP----GKPPDPGRPTQPGLSKSRTTDTFRSEQKLPG 123
                                                                                                                                                                                                                                                                                                                           124 RSPSTISLKESKSRTDFKEEYKSSMMPGFFSDVNPLSAVSSVVNKFNPFDLISDSBASQE 183
                                                                                                                                         74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LKSRPGTASSSGGSIARGAPVIV-----PBLGKPROSPLTYEDHGAPFAGHLPRGSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    IISQPGPGSESKD-----PKHIDPIQKKDEPKKAQPKGSP-----KPETKPVPKGSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PSSPOPTPKAAT-----APTATASKSPVPS--QOASPKKEPPSKQDS-PKA---LES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     686 APAAASEEAAFPPVVEDEEMEASGVSGNEEEMVEEAEALHASGNEVPRGECSGPATVNNS
                                                                                                                                                                                                                                                                                                                                                                                                          KSPAQPAGTGKSPAQPPAKTPGQQAGLEKTSSSQQPGPKSLAQTPGHGKFPLGPVKSPAQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ----NLOPESDAPQOPGSSPRGKSRSPAPPADKEAFAAEAQKLPGDPPCWTSGLPFPVPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -----IPEKANFNICTECOSTVCSLCGFNPNPHLTEIKEWLCINCOMQRALGGDLAAAI
                                                                                                                                                                                                                                                                                   ------SEATGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ---KSECTEEAE----EGPAKGKDAE--AAEATAEGAL--KAEKKEGGSGRATT----A
                                                                                                                                                                                                                                                                                                                                                                              PTPPPAPPSPSAPPVVPKEEKEETAAAPPVEEGEEQKPPAABELAVDTGKAEEPV---
                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             741;
Length 5085;
                                                                                                                                                                                           SDTESIPSPHTEAAKDTGONGPKPPATLGADGPPPGPPTPP----RRTS--
                                                 Indels
  Query Match 3.5%; Score 462.5; DB 1; Best Local Similarity 19.3%; Pred. No. 1.1e-07; Matches 402; Conservative 241; Mismatches 700;
```

```
2075
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1674
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1982
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1324
                                             -----AKSPHSTVPEHHPH----PISPYEHLLRGVSGVDLYRSHIPLAFDPTSIPR 1639
                                                                                                       1640 GIP-----LDAAAAYYLPRHLAPNPTYPHL----YPPYLIRGYPDTAALENR- 1682
                                                                                                                                                                  -----E 1717
                                                                                                                                                                                                                             1718 SSLALN----YAAGPRGIIDLSQVPHLPVLVPPTPGTPATAMDRLAYLPTAPQPFSSRH 1772
                                                                                                                                                                                                                                                                                       1773 SSSPLSPGGPTHLTKPTTTSSSERERDRDRERDRDREREKSILTSTTTVEHAPIWRPGTE 1832
                                                                                                                                                                                                                                                                                                                                                                     SKPT-VLRSTSTSSPVR--PAATFPPATHCPLGGTLDGVYPTLMEPVLLPKEAPRVARPE 1949
                                                                                                                                                                                                                                                                                                                                                                                                                               2020
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1325 IPSLIPSDEKDLLK------ĞLKKDSFSQES----SPSSPSDLAKLESTVLSIL 1368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SETLDITISBEBIKESQEKKVSPKKDSEQGFPSRK----EHKEKPELV-----DDLSPRR 1454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ------YDSVEDSSESENSPVVRRKRRTSIGSSSSDEYKOEDSQGSGEEED 1501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1502 FIRKQIIEMSADEDASGSEDEEFIRSQLKEISGVGESQKREEAKGKGKGVAGKHRLLTRK 1561
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1562 SSTSFDDDAGRRHSWHDEDDETFDESPELKFRETKSQESEELV---VAGGGGLRRFKTIE 1618
TPSGTRPTAGQAAPPSQQPPKPQBQSRRPSLNLGGITDAPKSQ----PTTPQETVTGKLF
                                                                                                                                                                                     PAHKPDKTTKPKPACPLCRTELN--LGSQEPPNFNTCTECKNQVCNLCGFNPTPHLTEIQ
                                                                                                                                                                                                                                               1054 EWLCINCOTORAISGOLG--DMGK-----KASPMP----
                                                                                                                                                                                                                                                                                                          936 QLPPAPAKATAVKKEAKAAAENLESKPEQAPTAKKTEKDKKPPPAKVCKPPPSEPEKAV
                                                                                                                                                                                                                                                                                                                                                   QSSGSSGSSGGGGGSSSRPASHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGIITAVEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ------AKPPARSGLEPASSPSKGSEPR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PLVPPVS-----GHATIARTPAKNLAPHHASPDPPAPPASASD-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1265 AVAQPQAEGSSKDGQGERSKEKTEKEEDKSDTSSSQQPKSPQGLSDTGYSSDGISGSLGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ----PHREKTQSKPFSIQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDKGLPKHLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ELDKSHLEGELRPKQPGPVKLGGEAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQRVVTL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EAQASTLVGEKAEKKTQPQKISPEK-----PQDQ------QKTQTA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2136 AQHISEVITQDYTRHHPQQLSAPLPAPLYSFPGASCPVLDLRRPPSDLYLPPPDHGAPAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GSPHSEGGKRSPEPNKTSVLGGGEDGIEPVSPPEGMTEPGHSRSAVYPLLYRDGEQTEPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -----RMGSKSPGNTSQPPAFFSKLTESNSAMVKSKKQE------INKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ---PEYNI----SQPGTEIFNMPAITGTGLMTYRSQA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2332 VQEHASTNMGLEAIIRKALMGKYDQWEESPPLSANAF------NPLNASASLP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAMP--ITAADGRSDHTLTSPGGGKAKVSGRPSSRKAKSPAPG 2420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GTSPTSVSSLDEDSDSSPSHKKGESK-----QQRKARHRSHG 1711
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RPRADTGHAFL-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LINT-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Š
                                                                                                                                                                                                966
                                                                                                                                                                                                                                                                                                                                                   1833
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2076
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1619
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1675
                 828
                                                                                                                                                                  1683
                                                                                                                                                                                                                                                                                                                     1090
                                                                                                                                                                                                                                                                                                                                                                                                             1893
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1950
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1983
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2256
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2294
                                             1593
                                                                                                                                                                                                                                                                                                                                                                                1123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2021
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                    ઠ
                                                                                                                              윱
                                                                                                                                                                 δ
                                                                                                                                                                                     a
                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                      Š
                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                  ਨੇ
                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ક
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ઠે
```

```
Pukuyama R., Rapoport S.I.;

Tukuyama R., Rapoport S.I.;

Thukuyama R., Rapoport S.I.;

Thukuyam
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Chen Z.C., Fadiel A., Naftolin F.; "Identification of a novel protein (P80) in ovarian carcinoma cells."; Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE-97001161; PubMed=8812494;
Fink J.K., Jones S.M., Esposito C., Wilkowski J.;
"Human microtubule-associated protein la (WAPIA) gene: genomic
organization, DNA sequence, and developmental- and tissue-specific
                 MAPA HUMAN STANDARD; PRT; 2805 AA. P78559; 095643; Q12973; Q15882; Q9UJT4; [15-UIJ-1998 (Rel. 36, Created) [16-OCT-2001 (Rel. 40, Last sequence update) [16-OCT-2001 (Rel. 40, Last annotation update) Microtubule-associated protein 1A (MAP 1A) (Proliferation-related protein protein P80) [Contains: MAPI light chain LC2].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ohtani K., Rutherford T., Sakamoto H., Naftolin F.;
"Microtubule associated protein 1A (MAP1A) in human brain sequence and physiological role.";
Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                              Chordata; Craniata; Vertebrata; Butele
Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TISSUE=Fetal muscle;
Chiannilkulchai N., Pasturaud P., Richard I., Auffray
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TISSUE=Brain;
MEDLINE=95356255; PubMed=7629894;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE OF 1607-1883 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL, U38291; AAB41132.1; -.
EMBL; U38292; AAB41133.1; -.
EMBL; AF200415; AAF08305.2; -.
EMBL; U80458; AAD00355.1; -.
                                                                                                                                                                                                                                                                              Eukaryota; Metazoa; Chordata;
Mammalia; Eutheria; Primates;
NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE OF 134-419 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE OF 78-1687 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE OF 1-1825 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Genomics 35:577-585(1996)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TISSUE=Ovarian carcinoma;
                                                                                                                                                                                                                                                       sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Beckmann J.S.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TISSUE=Brain;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              expression.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2
```

1361 EPKOPSULOGKOKTLEHKEVVEPKOTAIYQKOEALHYKNEAVKQOPKALEGKGRDLE 254 LPLYNQPSDTRGYHENIKINGAMEKKLILYFKRENHHARKQWKQKCQRYDQIMEAL 1417	893 ABKREGGGGRAFTAKSGGAPQDSDSSATCSADEVDBAEGGDKNRLLSPR 941 2000 LSTKEAAAGRNTSAEKELSSPISPKSLQSDTPTFSYAALAGPTVPPRXEPGPSME 20S4 942 PSLLTPTGDPRANASPQKPLDLKQLKQRAAAIP
	8 6 8 6 8 6 8 6 8 6 8 6 8 6 8 6 8 8 8 8
PRESENT EARLY STATUS CAMPTION.11	Query Match 3.4%; Score 451.5; DB 1; Length 2805; Best Local Similarity 21.8%; Pred. No. 1.4e-07; Matches 427; Conservative 199; Mismatches 732; Indels 599; Gaps 97; Qy 18 RYPPHSLS

us-09-522-753-5.rsp

```
----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      +
2397
                                                                                                                                                                                                                                                                                                                                                                                                        2692
                                                                                                                                                                                                                                                                                                                                                                                                                                1679
                                   2311 PASLDLALAPAPSLPGDMGDGILPCHLECSEAATEKPSPFQVPSEDCAANGPTETSPNPP 2370
                                                                                                                                                                                                                                                                      DTGASTTGSKK----HDVRSLIGSPGRTFPPVHPLDVMADARAL------ERACYEESL 1508
                                                                                                                                                                                                                                                                                                 2601
                                                                                                                                                                                                                                                                                                                                                    2633
                                                                                                                                                                                                                                                                                                                                                                             1624
                                                                                                                                                                  GSITQG----IPRSYVEAQEDYLRREAKLLKR------EGTPPPPPPSRDLTEAYKT 1399
                                                                                                                                                                                                                    1400 QALGPLKLKPAHEGLVATVKEAGRSIHEIPREELRHTPELPLAPRPLKEGSITQGTPLKY 1459
                                                                                                                                                                                                                                                                                                                                                                                                                                                          ----ANGEPSRAVL 2737
                                                                                      ---AWPEGA
                                                                                                                                                                                                                                    SDSG-----SSQSDSDVPPB----TEECPSITAEAALDSDEDGDFLPV
            PESLGVPTAQEASVLRGTALGSVP-----GGSITKGIPSTRVPSDSAITYRGSITHGTPA
                                                            DVLYKGTITRIIGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGR
                                                                                                              1296 SSSGPPHETAAPKRTYDWMEGRVGRAISSASIE---GLMGRAIPPERHSPHHLKEQHHIR
                                                                                                                                       2398 ERSSRPDTXLSPEQPV-CPXGGSGGPPSSASPEVEAGPQGCXTEPRPH------R
                                                                                                                                                                                          2446 GELSPSFLNPPLPPSIDD--RDLSTEEVRLVGRGGRRRVGGPGTTGXPXPVTDETPPTSA
                                                                                                                                                                                                                                                                                              2543 DKXGGVSGTHHPRPGHDPPPLPQPDXRPSPP-RPDVCMADPEGLSSESGRXERLRXKEKV
                                                                                                                                                                                                                                                                                                                         KSRPGTASSSGGSIARGAPVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSPVTMREPTPR
                                                                                                                                                                                                                                                                                                                                                  OGRVG-----RRAP-----GKDK--PV-----SPXRRLXLRGK----RSPTPG
                                                                                                                                                                                                                                                                                                                                                                             1569 LQEGSLSSSKASQDRKLTS--TPREIAKSPHSTVPEHHPHPI--SPYEHLLRGVSGVDLY
                                                                                                                                                                                                                                                                                                                                                                                                     2634 KGSXDRVSRXPXRSRSXTSQVTPAE-EKDGHSPMSKGLVNGLKAGPXALSSKGSSGAPVY
                                                                                                                                                                                                                                                                                                                                                                                                                                1625 RSHIPLAFDPTS-IPRGIPLD----AAAAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              28-FEB-2003 (Rel. 41, Created)
28-FEB-2003 (Rel. 41, Last sequence update)
10-OCT-2003 (Rel. 42, Last annotation update)
10-OCT-2003 (Rel. 42, Last annotation update)
Nuclear receptor coactivator 6 (Amplified in breast cancer-3 protein)
(Cancer-amplified transcriptional coactivator ASC-2) (Activating signal cointegrator-2) (ASC-2) (Peroxisome proliferator-activated receptor-interacting protein) (Parkinteracting protein) (Nuclear receptor-activating protein, 250 kba) (Nuclear receptor coactivator RAP250) (NRC) (Thyroid hormone receptor binding protein).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           "Isolation and characterization of peroxisome proliferator-activated receptor (PPAR) interacting protein (PRIP) as a coactivator for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Chordata, Craniata, Vertebrata, Buteleostomi, Rodentia, Sciurognathi, Muridae, Murinae, Mus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE=20250907; PubMed=10788465;
Zhu Y.-J., Kan L., Qi C., Kanwar Y.S., Yeldandi A.V., Rao M.S.,
Reddy J.K.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND INTERACTION WITH PPARA; PPARG; RARA; RXRA; ESR1; ESR2 AND THRB.
TISSUE=Liver;
                                                                                       ------AWERG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                        2693 ---VDLAYIPNHWSGKTADLDFFRRVRASYYVVSGNDP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ------ENRQ---TIINDYITS-----QQMH 1696
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2738 DALLEGKAQWGENLÓVKVTLÍPTHDŤEVTREWYQÓTH 2774
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2067 AA.
                                                                                     ----XPAPAKAENEEAAAXP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Biol. Chem. 275:13510-13516(2000)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A. (ISOFORM 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Bukaryota; Metazoa; Chordata;
Mammalia; Butheria; Rodentia;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Mus musculus (Mouse)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCO6 MOUSE (
                                                            1236
                                                                                     2371
                                                                                                                                                                  1353
                                                                                                                                                                                                                                                                      1460
                                                                                                                                                                                                                                                                                                                         1509
                                                                                                                                                                                                                                              2504
                                                                                                                                                                                                                                                                                                                                                   2602
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1680
                                                                                  a
                                                                                                              ઠે
                                                                                                                                       셤
                                                                                                                                                               ઠે
                                                                                                                                                                                       셤
                                                                                                                                                                                                                  ሯ
                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                     ઠે
                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                           ò
                                                                                                                                                                                                                                                                                                                        ઠે
                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                     음
                                                                                                                                                                                                                                                                                                                                                                                                                                ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 è
```

```
MEDLINE=21638469; PubMed=11704680;

AM MEDLINE=21638469; PubMed=11704680;

A Jung D.-J., Na S.-Y., Na D.S., Lee J.W.;

A Jung D.-J., Na S.-Y., Na D.S., Lee J.W.;

A Jung D.-J., Na S.-Y., Na S.-Y.,

A Jung D.-J., Na S.-Y., Na S.-Y.,

BAT J. Biol. Chem. 277:1229-1234 (2002).

L. J. Biol. Chem. 277:1229-1234 (2002).

L. FUNCTION: Nuclear receptor coactivates expression in an agonist-correct dependent fashion. Coactivates expression in an agonist-correct dependent manner. Involved in the coactivation of different nuclear receptors, such as for steroids (GR and ERS), and prostanoids (RARS and RXRS), thyroid hormone (TRS), vitamin D3 (VDR) and prostanoids (PARS), probably functions as a general coactivator, rather than just a nuclear receptor coactivator. May coactivator, rather than just a nuclear receptor coactivator. May coactivate expression via a remodeling of chromatin and its interaction with histone acetyltransferase proteins. Involved in the coactivator and embryonic development.

C. SUBUNIT: Monomer and homodimer. Interacts in vitro with the basal transactivation function. Interacts with NCOA1, CRSPBP, and WIRH, the methyltransferase proteins EP30 and CRSPBP, and wind which contains ASC-2/NCOA6 complex (ASCOM), complex of proper and homedime expension and with methyltransferase proteins EP30 and CRSPBP, and wind in nuclear such RNPILLI (By similarity).

C. Interacts with RNPC2. Belongs to the ASC-2/NCOA6 complex (ASCOM), complex of the interior are and health and heal
                       Astraubberg R.D., Feingold E.A., Grouse L.H., Derge J.G.,
Riausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altachul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Mang J., Haibh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Blatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Millahy S.J.,
Raha S.S., Morley K.C., Hale S., Garcia A.M., Gabbs R.A.,
Nillalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Nillalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Mitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Blakesley R.W., Touchman J.W., Schalska U., Smailus D.E.,
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
Generation and initial analysis of more than 15,000 full-length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gustafsson J.-A.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE OF 786-1142 FROM N.A. (ISOFORM 1), INTERACTION WITH PPARA; PPARG; ESR1; ESR2; THRA AND THRB, AND MUTAGENESIS OF LEU-891 AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           'Note-Acts as a dominant negative repressor;
-!- TISSUE SPECIFICITY: Widely expressed. High expression in testis and weak expression in small intestine.
-!- DEVELOPMENTAL STAGE: Expressed at E9 in placenta and at weaker level in uterus. High expression in neural tube and in CNS throughout development. High expression in sensory ganglia and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RBQ-3/ RBBP5, alpha- and beta-tubulins, the trithorax group proteins Mil2 and Mil3, and ASH2/ASCL2 (By similarity). SUBCELLULAR LOCATION: Nuclear.
ALTERNATIVE PRODUCTS:
Event-Alternative splicing; Named isoforms=2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Caira F., Antonson P., Pelto-Huikko M., Treuter E., Gustafs:
"Cloning and characterization of RAP250, a nuclear receptor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Isold=09JL19-2; Sequence=VSP 003410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IsoId=Q9JL19-1; Sequence=Displayed;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        coactivator.";
J. Biol. Chem. 275:5308-5317(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TISSUE=Embryo;
MEDLINE=20148724; PubMed=10681503;
MEDLINE=22388257; PubMed=12477932;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Name=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Name=2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LEU-894.
```

```
1593 AKSPHS--TVPEHHPHPISPY-EHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLDAAAAY 1649
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PGTASSSGGSIARGAPVI-----VPELGKPRQ----SPLTYEDHG----APFAGHLPR 1556
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RGSITQGIPRSYVEAQEDYLRREAKLL---KREGTPPPP----PPSRDLTE-----AYK 1398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TQGTPLKYDTGASTTGSKKHDVRSLIGSPGRTFPPVHPLDVMADARALERACYEBSLKSR 1511
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1004 NIQPESDAPQQ-PGSSPRGKSRSPAPPADKEAFAAEAQKLPGDPPCWTSGLPFPVPPREV 1062
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SVLERQIGAISQGMSVQLHVPYSEHAKAPVGPVTWGLPLPMDPKKLAPFSGVKQEQLSPR 1175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1235 ADVLYKGTITRIIGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMS-VTQCS--K 1291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1292 EDGRSSSGPPHETAAPKRTYDMMEGRVGRAISSASIEGLMGRAIPPERHSPHHLKEQHHI 1351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PUTTANNDVNISQMMPDVSMQQASMVPPHVQSMQGNSASGSHFSGHGVSFNAPFGG-APN 866
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    700 NOMMGPOGQVLLQQNPMIEQIMTNOMQGNKAQFNSQNQSNVMPGPAQIMRGPTPNMQGNM 759
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              643 QGTLNPQNPMILSRAQ----LMPQGQMMVNAQNQNLGPSPQRMTPPKQMLPQQGPQMMAPH 699
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GPATVNNSSDTESIPSPHTEAAKDTGQNGPKPPATLGADGPPPGPPTPPRRTSRAPIEPT 797
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GSOMSCGONPGFPVNKDVTLTSPLLVNLLQSDISAGHFGVNNKONNTNANKPKKKKPPRK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               760 VQFTGQMSGQMLPQQGPVNNSPSQVMGIQGQVLRPPGPSPHMAQQH-----TD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        483 MV---QGINVPENEWWQQQPENQ-----GPQSLHPGLGGMPKRLPP--GFSAGQANPN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GQAGPPESLGVPTAQEASVLRGTALGSVP-GGSITKGIPSTRVPSDSAITYRGSITHGTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          582 NWM------QPSLM-----QPSLM-----GIHGNINNQQAGSSGVPQVTLGN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ----SQLMGMHQQIVP----SQGQMAQQ---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      532 FMOGQVPSTTAATPGNSGALOLOANONVQHAGGQGAGPPONOM------OVSHGPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TQALGP-----LKLKPAHEGLVATVKEAGRSIHBIPREE--LRHTPELPLAPRPLKEGSI
                                                                                                                                                                                                                                                                                                                                                                                                                                     LLTPTGDPRANASPQKPLDLKQLKQRAAAIPPIQVTKVHEPPREDAAPTKPAPPAPPPQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            .063 IKASPHAPDPSAFSY-APPGHPLPLGLHDTARPVLPRPPTISNP-----PPLISSAKHP
                                                                                                                                                                                                                   210 PASQSDAMDPLLSGLHIQQQSHPSGSLPPAHHSMQPVPVNRQMNPANFPQLQQQQQQQQ
                                                                                                                                                                                                                                                              832 AAAPPVEEGEEÇKPPAAEELAVDTGKAEEPVKSECTEEAEEG------PAKGKDAEAAE
                                                                                                                                                                                                                                                                                         270 QQQQQQQQQQQQQQQQQQQQQQQQQQQQQT~---RPLQQH--QQQPQGIRPQFTAPTQVPPGWN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              383 GSQQASQAHTNFPQMSNPGQFTAP-----QMKGLQGGP----SRVPTPLQQPHL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TNKSP-ASSPESFQQGSPASSPT---VNQTQQMGPRPPQ-NNPLSQGFQQPVSSPGRNP
                                                                                                                                                                                                                                                                                                                                                 885 ATARGALKAEKKEGGSGRATTAKS-SGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPS
                                          114 SNNQQLRDLGILSVQ---IEGEGAINLALGQNRSQDVRMNGPVASGNSV-RMEAGFPMAS
                                                                                                                                GPGLIRMTS-----PAAVMTPOGGNMSSSMM---APGPNPELOPR-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GS-------PVTMREP-TPRLQEGSLSSSKASQDRKLTST-
                                                                                                                                                                                                                                                                                                                                                                                                 324 QLPSGALQPPPAQGSLGTMTTNQGWKKAPLPSPMQAQ-----
                                                                                                                                                                                    PASEA-----TGAPTPPPAPPSPSAPP
             700 VEDEEMEASGVSGNEEEMVEEAEALH-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           367 LAT-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        613 МОĞОРООĞРР-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1557
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        808
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         867
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1512
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1176
                                                                                                                                                                                                                                                                                                                                                                                                                                                     944
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  428
                                                                                                                                             170
                                                                                                                                                                                        798
                                                                                                     738
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                     유
                                                                                                                                                                                                                                                                                                                                                                                                                                                     8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                               ò
                                                   셤
                                                                                                     ð
                                                                                                                                           원
                                                                                                                                                                                        ò
                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                   this SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NSEEAITPQQSAELASMELNESSRWTEEEMETAKKGLLEHGRNWSAIARWVGSKTVSQCK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             651 NFYFNYKKRQNLDEIL------QQHKLKMEKER--NARRKKKKAPAAASEEAAFPPV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ---FKWK----LDAILKNVPNLLHMESSKLKVQKVEPWNSVRVTFNIPREAAERLWILAQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NFEDIYTSLCSSTMGDSEVEFDSGL--EDDDTKGDSLLEDSTIFVAFKGNIDDKD----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CREBBP-BINDING REGION (BY SIMILARITY).
TBP/GTF2A-BINDING REGION (BY SIMILARITY).
NCOA1-BINDING REGION (BY SIMILARITY).
NCOACIP-BINDING REGION (BY SIMILARITY).
EP300/CRSF3-BINDING REGION
retina from Ell. In the alimentary tract and olfactory epithelium expression was seen from El3. Strong expression present in liver and kidney, from El1 and El3 respectively, and then expression decreased al later stages of development. Moderate expression in lung from El3, while it decreases during postnatal life. Strong expression in thymus from El5 onwards, and in spleen from El7 and during early postnatal life, then, the expression decreases. DOMAIN: Contains two leu-Xaa-Xaa-Leu-Leu (LXXLL) motifs. Only motif is essential for the association with nuclear receptors. PTM: Phosphorylated (By similarity).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              II promoter; IDA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LXXLL MOTIF 1.

LXXLL MOTIF 2.

Missing (in isoform 2).

/FTG=VSP 003410.

LVNL->AVNĀ: ABOLISHES INTERACTION WITH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DAMBLE STATES ST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     731;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 2067;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Transcription regulation; Activator; Nuclear protein; Repeat;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                C855F8777167AD48 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    G -> S (IN REF. 2).
W -> R (IN REF. 2).
M -> I (IN REF. 2).
Q -> QQ (IN REF. 2).
P -> L (IN REF. 3).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3.3%; Score 439.5; DB 1;
19.4%; Pred. No. 2.6e-07;
ive 295; Mismatches 803;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NUCLEAR RECEPTORS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (IN REF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SE -> RS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SER-RICH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative 295;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          194
290
1014
1142
\; 219663 M·
                                                                                                                                                                                                                                                                                                                                                                                                                                            BC031113; AAH31113.1; -. AF135169; AAF35973.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; AF216186; AAF35860.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1592
895
1499
2067
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1314
931
2067
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  894
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Alternative splicing.
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1014
1141
2067
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             39
109
194
290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    891
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              440;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       591
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /ARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local
Matches 44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MUTAGEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
```

531

581

482

943

--- PVVPKEEKEET

-----LQARPS

1451

us-09-522-753-5.rsp

```
The
                                                                                                                                                                                                                                                                                                                                                                                                                                                            \Xi
                                                                                                                                                                                                                                                                                                                                                                                                    2018
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1477
                                                                                                                                                                                                                                                                                                                1304
                                                                                                                                                                                                                                                                                                                                                                        1364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2105
                                                                                                                                                                                                                                                                                                                                                                                                                               1421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2393
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | | | | | | | | | SANRRSPVSSSKGKGKVDKIGQILLITKACKKVTGSLEKGEEQYGADGETEGPGLEITTP 1865
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1663 FINSSIIQVMKG-------SQPST-IPATPLTINSGLMPPSVAVVGPL 1703
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2458
  -- GYPDTAALENRQTIINDYITSQQ
                1245 ERLNASIAGLFPPQINIPLPPRPNLNRGFDQQGLNPTTLKAIGQAPSNLTITNPPNFAAP
                                                                                                                                                                                                                                                                                                                                                                                                                    STIMLONMELPRNV---LVGPTPLANPPLPGSFPNNTGLNPQNPTVPVPAMGTVLEDNKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SVNIPQDSDCQNAQGRKEQVNTELKVVPTQEAKMAV----PEDQSKKDGQPLDPNKLPSV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HIPQNI------KF----SSAPVTPNV--PSSSPAPNIQTGRPLVLSSRATPVQ
                                                       1695 MHHNTATAMAQRADMLRG-LSPRESSLALNYAAGPRGIIDLSQVPHLPVLVPPTP----
                                                                                                               -- PGGPTHLTKP
                                                                                                                                          PMSVNTPMGSNSRKMVYQENPQ - - - - NSSSSPLGEMSSLPEASGSEVPSVAGGPNNMPSH
                                                                                                                                                                    TITISSSERERDRDRERERERSILISTITIVEHAPIWRPGTEQSSGSSGGGGGGSS
                                                                                                                                                                                               1157 LVVSQNQ-------LMMTGP-----KPGPSPLSATQGATPQQPPVN
                                                                                                                                                                                                                            SRPASHSHAHQHSP-ISPRTQDALQQRPSVLHNTGMKGIITAVEPSKPTVLRSTSTS-SP
                                                                                                                                                                                                                                                     SLPSSHGH---HFPNVAAPTOTS---RPKTPNRASPRPYYPQTPNNRPPSTEPSEISLSP
                                                                                                                                                                                                                                                                                    1907 VR----PAATFPPATHCPL-----GGTLDGVYPTLMEPV-----LLPKEAPRVARP
                                                                                                                                                                                                                                                                                                                                           ERPRADT------GHAFLAKPPARSGLEPASS----PSKGSE-----
                                                                                                                                                                                                                                                                                                                                                                     QAHKLDSVVVNSGKQSNPGTTKRASPSNSRRSSPGSSRKTTPSPGRQNSKAPKLTLASQT
                                                                                                                                                                                                                                                                                                                                                                                                   -----PRPLVPPVSGHATIARTPAKNLAPHHA--SPDPPAPASA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                         -----SD-----PHREKTQSKPFSIQELELRSLGYHGSSYSPEGVEPVSPVSSPSL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----PVKLGGE----AAHLPH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LRPLPESOPSSSPLLQTAPGVKGHQRVVTLAQ-HISEVITQDYTRHHPQQLSAPLPAPLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LODPPSKEPSTS---LSSP--HSSEPCSTLARSELSEVSSN------AAPSIPPVM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SPPGASCPVLDLRRPPSDLYLPPPDHGAPARGSPHSEGGKRSPEPNKTSVLGGGEDGIEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2225 VSPPEGMTE--PGHSRSAVYPLLYRDGEQTEPSRMGSK-----SPGNTSQPPAFFSKL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2276 TESNSAMVKSKKQEINKKLNTHNRNEPEYNISOPGTEIFNMPAITGTGLMTYRSQAVQE-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HASTNMGLEAIIRKALMCKYDQWEESPPLSANAFNPLNASA-SLPAAMPITAADGRSDHT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2394 LTSPGGGGKAKVSGRPSSRKAKSPAPGLAS------GDRPPS----VSSV--HSEG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----STPFPYNPLIMRLQAGVMA---SPPPPGLPAGSGP 2490
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1866 GLMGTEQCSTELDSKTPTPSAPTLLKMTSSPMAPSSTSTGPILPGGALP 1914
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2065 THDKGL-----PKHLEELDKSHLEGELRPKQPG-
                                                                                                               ----GIPATAMDRLAYLPTAPOPFSSRHSSSPLS--
 1650 YLP-----RHLAPNPTYPHLYP-PYLIR-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2438 -DCNRRTPLTNRVWEDRPSSAG
                                                                                                                                          1101
                                                                                                                                                                                                                                                        1191
                                                                                                                                                                                                                                                                                                                                           1949
                                                                                                                                                                                                                                                                                                                                                                       1305
                                                                                                                                                                                                                                                                                                                                                                                                                            1365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1538
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1609
                                                                                   1045
                                                                                                               1749
                                                                                                                                                                                                                            1849
                                                                                                                                                                                                                                                                                                                                                                                                                                                         2019
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1583
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1704
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1806
                                                                                                                                                                      1789
                                                                                                                                                                                                                                                                                                                                                                                                   1981
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2459
                           셤
                                                                              셤
                                                                                                            ò
                                                                                                                                     셤
                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                       셤
ò
                                                       ሯ
                                                                                                                                                                 ઠે
                                                                                                                                                                                                                         ઠે
                                                                                                                                                                                                                                                                              ઠ
                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ሯ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           6 6 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
```

```
MEDLINE=22737999; PubMed=12853948;

MEDLINE=22737999; PubMed=12853948;

MEDLINE=22737999; PubMed=12853948;

MEDLINE=22737999; PubMed=12853948;

MEDLINE=22737999; PubMed=12853948;

MAILIER L. W., Fulton R.S., Fulton L.A., Graves T.A., Pepin K.H.,

Magner-McPherson C., Layman D., Mans J., Schaler M.E.,

RA Wylie K., Sekhon M., Becker M.C., O'Laughlin M.D., Schaler M.E.,

RA Sun H., Edwards J., Bradshaw-Cordum H., Ali J., Andrews S., Isak A.,

Sun H., Edwards J., Du F., Lamar B., Courtney L., Kalicki J.,

RA Sun H., Mayuyen C., Du F., Lamar B., Courtney L., Kalicki J.,

RA Colowicz-Reilly A., Leonard S., Rohlfing T., Rock S.M.,

Tin-Wollam A.-M., Abbott A., Minx P., Maupin R., Strowmatt C.,

Andreal B., Millar N., Johnson D., Murray J., Woessner J.P.,

Mendl M.C., Yang S.-P., Schultz B.R., Wallis J.W., Spieth J.,

Hickenbotham M.T., Eldred J., Milliams D., Bedell J.A., Mardis B.R.,

Alickenbotham M.T., Eldred J., Milliams D., Bedell J.A., Mardis E.R.,

Gillett W., Zhou Y., Jaman R.A., Raymond C., Haugen B.,

Simms B., Levy R., Clendenning J., Kaul R., Kent W.J., Furey T.S.,

Bailey J.A., Portnoy M.E., Torrents D., Chinhalla A.T., Gish W.R.,

Reddy S.R., McPherson J.D., Olson M.V., Eichler B.E., Green B.D.,

Wature 424:157-164(2003).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Nagase T., Kikuno R., Ishikawa K.-I., Hirosawa M., Ohara O., "Prediction of the coding sequences of unidentified human genes. XVII. The complete sequences of 100 new cDNA clones from brain which code
                                                                                                                  homolog (Histone-3) (EC 2.1.1.43)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        to ALR and associated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
                                                                                                                                                                                                                                                                                                                                TISSUE=Fetal thymus;
MEDLINE=21888622; PubMed=11891048;
Ruault M., Brun M.-E., Ventura M., Roizes G., De Sario A.;
"MLLJ3, a new human member of the TRX/MLL gene family, maps to 7q36,
chromosome region frequently deleted in myeloid leukaemia.";
Gene 284:73-81(2002).
                                                                                                                                                                                                                            Craniata; Vertebrata; Euteleostomi;
Catarrhini; Hominidae; Homo.
MLL3 HUMAN STANDARD; PRT; 4911 AA. OBNEZ4; QBNDC02; QBNDF6; Q9H94; Q9NR13; Q9P222; Q9UDR7; 10-OCT-2003 (Rel. 42, Created) 11-OCT-2003 (Rel. 42, Last sequence update) 15-MAR-2004 (Rel. 43, Last sannotation update) Myeloid/lymphoid or mixed-lineage leukemia protein 3 homolog [Homologue to ALR protein). Myeloid/Lymphoid or Mixed-lineage leukemia protein 3 homolog Myeloid/Lymphoid or Mixed-lineage leukemia protein 3 homolog Homologue to ALR protein).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TISSUE=Cervical carcinoma;
MEDLINE=21574953; PubMed=11718452;
Tan Y.C., Chow V.T.;
Tan Y.C., Chow W.T.;
Tan Y.C., Chow W.T.;
To Mile I man HALR (Mile I) gene encodes a protein homologous to ALL-1 involved in leukemia, and maps to chromosome 7q36 with leukemia and developmental defects.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       [5]
SEQUENCE OF 3193-3865 AND 4460-4911 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE OF 556-3865 FROM N.A. (ISOFORM 1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE=20277482; PubMed=10819331;
                                                                                                                                                                                                                                                                                                              (ISOFORM 1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A. (ISOFORM 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    vitro.";
                                                                                                                                                                                                                          Eukaryota; Metazoa; Chordata;
Mammalia; Eutheria; Primates;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    large proteins in
                                                                                                                                                                                                        Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                              FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   rissum=Placenta;
                                                                                                                                                                                                                                                                    NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                            SEQUENCE
```

```
Takeshashi W. Catha Y. Tabila S., Wartakawa K., Good Y. Takiguchi S.,
Yemanor C. ; Wakkmatu M., Nakhami Y., Ishii S., Kawai Y., Saito K.,
Yemanor C. ; Wakkmatu T.,
Windoniya K., 1994/mail T.;
Windoniya W., 1994/mail T.;
Windoniya K., 1994/mail T.;
Windoniya K., 1994/mail T.;
Windoniya W., 1994/mail T.;
Windon
```

```
.---G 1567
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                91 YL-----PELGKSEMEFIESKRPRLELLPDPLLRPSPLLATGOPAGSEDLTKDRSLTGK 144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  145 LEPVSPPSPHTDPELELVPPRLSKEELIQNMDRVDREITMVEQQISKLKKKQQQLEEEA 204
                                                                                                                                                         34 HTDVGLLEYQHHSRDYASHLSPGSIIQP---QRRRPSLLSEFQPGNERSQELHLRPESHS
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
3.3%; Score 439; DB 1; Length 4911;
Best Local Similarity 18.3%; Pred. No. 6e-07;
Matches 556; Conservative 313; Mismatches 1010; Indels 1152;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POST-SET.
COLLED COLL (POTENTIAL).
A. T HOOK (BY SIMILARITY).
GLN-RICH.
PRO-RICH.
RSP-RICH.
                                                                                                                                                                                                                                                                                            DHHC-TYPE.
PHD-TYPE 3.
PHD-TYPE 4.
PHD-TYPE 5.
                                                                                                                                                                                                                                                                              RING-TYPE.
PHD-TYPE 2.
                                                    FYrich C.
FYrich N.
HMG 12 box.
PostSET.
                                                     InterPro; IPR003889; FYTich C.
InterPro; IPR003889; FYTich N.
InterPro; IPR000910; HMG 12_box.
InterPro; IPR001214; SET.
InterPro; IPR001214; SET.
InterPro; IPR001594; Znf_DHHC.
InterPro; IPR001965; Znf_PHD.
InterPro; IPR001965; Znf_PHD.
Pfam; PP00505; HMG_box; I.
Pfam; PP00628; PHD; 6.
Pfam; PP00628; PHD; 6.
AB040939; BAA96030.2; -. AK022687; BAB14179.1; -. AK075113; BAC11409.1; -. AL833924; CAD38780.1; -.
                                              InterPro; IPR000637; AT hook.
                              HGNC:13726; MLL3.
                                                                                                                                                                                                                                                                                                                                                                                            3272
3433
46
1796
2281
2630
2786
                                                                                                                                                                                                                                                                                               489
520
1010
1010
1139
4891
1136
672
1366
1787
3081
                                                                                                                                          SMART; SM00542; FYRC; 1
SMART; SM00541; FYRN; 1
                                                                                                                                                                                                                                                                                                                                     4895
92
644
11338
11734
3173
3391
3391
17134
17134
1834
2690
                                        - '883999
                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
DOMAIN
DNA BIND
                                                                                                                                                                                                                                                                                               ZN_FING
ZN_FING
ZN_FING
ZN_FING
ZN_FING
                                                                                                                                                                                                                                                                                  ZN FING
                                                                                                                                                                                                                                                                                                                                      DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                              DNA BIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                 Genew;
         EMBL;
EMBL;
EMBL;
임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
```

2330 SEGSFCASNSPWHSGGQCFSGVSQL; 1216RUPSDS-AIT 2388 REIILQQQQCKXIAGRQEKGSQDSPAVPI 248 GEDSPSRLDRG-REDSLPKGH 2445SPVAPPLGPRYAVFFKDQRGFYPPD 2445SPVAPPLGPRYAVFFKDQRGFYPPD 1298 SGPPHETAAPKRTYDMMEGR' 1299 LVPPQQIQGSGVSPQLRRSVSVDM 7499 LVPPQQIQGSGVSPQLRRSVSVDM 7499 LVPPQQIQGSGVSPQLRRSVSVDM		1706 RADMLEGLSPRESSLALMYAAGPRGIIDI 2986 GVQVNPCLIPQGST 1766 QPFSSRHSSPLSPGGSTHLTKPTTTSS 1766 QPFSSRHSSPLSPGGSTHLTKPTTTSS 1811 EKSILTSTTTVEHAPIWRPGTEQS 1811 EKSILTSTTTVEHAPIWRPGTEQS 1812 MQAMIRQRSEPFPNIDFDAITDFINKAI 1815SGSSGSGGGGGSSSRPASHSHAHQHE 1815SGSSGSGGGGGSSSRPASHSHAHQHE 1815SGSSGSGGGGGSSSRPASHSHAHQHE 1817 QQRPSVLHNTGMKGIITAVEPSI 1872 QQRPSVLHNTGMKGIITAVEPSI 1872 QQRPSVL
8 6 8 6 8 6 8 6 8 6 8	8 4 8 4 8 4 8 4 8 4 8 4 8 4 8 8 8 8 8 8	\$ 8 % 6 8 6 8 6 8 6 8 6 8 6 8 8 8 8 8 8 8
AKPPEPEKPUSPPPIESKHRSLVQIIYDENRKKABAAHRILEGLGPQVELPLYNOPSDTR : : : : : :	1721 QQUSIDPSSKIUSELKEDPLKQRESEHEQB	885 ATABGALKAEKKEGGSGRATTAKSSGAPQDSDSSATCSADEVDEAEGGDKNRILSPRPSL 944 2002 GTSDHFTKPSPRADVFQRIP-DSYARPLITPAPLDSGPGFKTP-MQPPPSS 2053 945 LTPTGDPRANASPOKPLDLKQLKQRAAAIPPIQVTKVHEPPREDAAPTKPAPPAPPPPQN 1004 2054 QDPYGSV-SQASRRLSVDPYERPALTPRPIDNFSHNQSNDPYSQPPTFPHPAVNES 2108 1005 LQPESDAPQQPGSSPRGKSRSPAPPADKEAFAAEAQKLPGDPPCWTSGLPPP 1056 2109 FAHPSRAFSQPGTISRTSQDPYSQPPGTPRPVUDSYSQSSGTARSNTDPYSQPFGTPRP 2168 1057VPPREVIKASPHAPDFSAFSYA-PPGHPLP-LGLHDTARPVLP 1097 2169 TTVDPYSQQPGTPSRTGYDLFVTPVTNQRHSDPYAHPPGTSRYSQPATP 2225 1098 RP
3 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	2 6 8 6 8 6 8 6 8 6 8	6 8 6 8 6 8 6 8 6 8 6

:|| | |: | h.pgpvptsgvtdtqntvnmaqadteklrqrqkl 2387 RVGRAISSASIEGLMGRAIPPERHSPHHLK-EQ 1348 --PRPLNNSQMNNPVGL---PQHFSPQSLPVQQ 2550 PRSYVEAQEDYLRREAKLLKREGTPPP--PPPS 1390 VKEAGRSIH--EIPREELRHT--PELPLAPRPL 1446 RSLIGSPGRIFPPVHPLDVMADARALE-RACYE 1505 TNDPNLDDLLRSGEFDIIAYTDPELDMGDKKSM 2772 LOEGS--LSSSKASQDRKLTST----- 1588 | | : | | | : | | EQENKTLVLSPN 2830 EHHP----- 1616 PDTAALENROTIINDYIT-SQOMHHNTATAMAQ 1705 -VSGHATIARTPA-----KNLAPHHASPDDP-- 2012 :||| : | || : || TYRGSITHGTPADV------LYKGTITRII 1247 : | : | | : | PHPGPLQHWQPENVNQAFTRPPPPYPGNIR--- 2444 -VIYEGKKGH--VLSYEGGMSVTQCSKEDGRSS 1297 DD-----PSVKELDV-KDLEGVEVKDLDD 2714 -----GAPVIV----PELGKPRQSPL 1541 AHSDLNDGEKTSLHPCDPDLFEKRTNRETAGPS 2890 RSHIPLA---FDPTSI-PRGIPLDAAAYYLPR 1653 DLSQVPHLPVLVPPTPGTPATAMDRLAYLPTAP 1765 ------QTGP 3015 SSERER-----1810 |||| QQNRERPLLLEEQPLLLQDLLDQERQEQQQQRQ 3071 HSPISP-----RTQDAL 1871 ISRPNPPNFGPGFVNDSQRKQYEEWLQETQQLL 3191 SKPTVLRSTSTSSPVRPAATFPPATHCPLGGTL 1925 :| akortakkagreppeedaeqlkhvte----- 3239 ADTGHAFLAK-----PPA-----RSGLE 1970 AELIEDYRIKQQQQCAMAPPTMMPSVQPQPPLI 3291 AKMVALKGINKVMAQNNLGMPPMVMSRFPFMGQ 3131

```
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                          2439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3804
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3805 IKDNKL---VEKQNPAEGLQT-----LGAQMQGGFGCGNQLPKTDGGSETKKQRSKRTQ 3855
                                                                                                                                                                                                                                           2147 YTRHHPQQLSAPLPAPLYSFPGASCPVLDLRRPPSDLYLPPPDHGAPARGSPHSEGGKRS 2206
                                                                                                                                                                                                                                                                1 : | : | | : | | 3664 | 3607 TRKKKRD--DDAESTKAPSTPHOSTPHISETTSTPAVSTPSELPQQADQESVBPV 3664
                                                   OPPIAQLPIKTCTPAPGTVSNANPQSGPPRVEFDDNNPFSESFQERERKERLREQQERQ 3411
                                                                                               | :|: |: |: | | | | RIQLMQEVDRQRALLQQRNEMEQHGWVGSBISSSRTSVSQIPFYSSDLPCDFMQPLGPLQQ 3471
PGATPPTMSQPTFPMVPQQLQHQQHTTVISGHTSPVRMPSLPGWQPNSAPAHLPLNPPRI 3351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAINBLISTOL N2;
MEDLINE-94150718; PubMed=7906398;
Milson R., Ainscough R., Anderson K., Baynes C., Berks M.,
Bonfield J., Burton J., Cornell M., Copsey T., Cooper J., Coulson A.,
Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
Johnston L., Jones M., Kerblaw J., Kirsten J., Laisster N.,
Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
Sims M., Smaldon N., Smith M., Sonnhammer E., Staden R.,
Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
                                                                                                                                                                                                                                                                                              2207 PEPNKTSVLGGGEDGIEPVSPPEGMT---EPGHSRSAVYPLLYRDGEQTEPSRMGSKSPG
                                                                                                                                                                                                                                                                                                                                                                        3665 GPSTPNMAAGQLCTELENKLP-----NSDFSQATPNQQTYANSEVDKLSMETPAKT--
                                                                                                                                                                                                                                                                                                                                                                                                    2323 GLMTYRSQAVQEHASTNWGLEAIIRKALMGKYDQWEESPPLSANAFNPLNASASLPAAMP
                                                                                                                                                                                                                                                                                                                                                                                                                                                          2383 ITAADGRSDHTLTSP---GGGGKAKVSGRPSSRKAKSPAPGLASGDRPPSVSSVHSBGDC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3753 VSSAQS-PPHSAGAPAAKGDSGNELLKHLLKNKKSSS----LLLNQKPE--GSICSEDDC
                                                                                                                                                                                                                                                                                                                                                  2264 NTSOPPAFFSKL-TESNSAMVKSKKOEINKKLNTHNRNEPEYNISOPGTEIFNMPAITGT
                                                                                                                                                                                                                                                                                                                                                                                                                                ------EBIKLEKAETESCPG----QEEPKLEEQNGSKVEGNA---VACP
                                                                                                                                   2062 PSLTHDKGLPKHLEEL----DKSHLEGELRP-------KQPGPVKLGGEAAHLP
                                                                                                                                                        3524 VGSPNFSSVKQCHGNLSGTSFQQSPVRPSFTPALPAAPPV-----ANSSLPCGOD
                                                                                 ---EGVEPVSPVSS
                                                                                                                                                                                        -----QPSSSPLLQTAPGVKGHQRVVTLAQHISEVITQD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Caenorhabditis elegans.
Bukaryota, Metazoa, Nematoda, Chromadorea, Rhabditida, Rhabditoidea,
Rhabditidae, Peloderinae, Caenorhabditis.
NCBI_TaxID=6239;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2440 NRRIPLINRVWEDRPSSAGSIPFPYNPLIMRLOAGVMASPPPPGLPAGS
                              ----APPASASDPHREKTQSKPFS---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           P34373; P34332;
01-FBB-1994 (Rel. 28, Created)
28-FBB-2003 (Rel. 41, Last sequence update)
28-FBB-2003 (Rel. 41, Last annotation update)
Hypochetical protein C1489.6 in chromosome III.
C1489.6/C1489.5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    --GPLAGPH---HAWDEEPKPLLCSOYETLS 2514
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RTGEKAAPRSKKRKKDEEEKQAMYSSTDTFT 3886
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ź
                                                                                  ---IQELELRSLGYHGSSYSP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STANDARD;
                                                                                                                                                                                              2105 HLRPLPES-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                           3574 STITHG--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CAEEL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2489
                                                                                                                                                                                                                                                                                                                                                                                                                                      3716
                                                                                                              3412
                                                                                      2033
                                                                                                                                                                   3472
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                              8
                                                                                                                                                                                        ઠે
                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                     임
                                                                                                                                                                                                                                                                                                    Š
                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                  õ
                                                                                                             g
                                                                                                                                         ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              270 IKINQAMRKKLI-----LYPKRRNHARKQWKQKFCQRYDQLMEALEKKVERIENN 319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LKVNSGLDLLLIFCSFPCLFDVLVFTFKELARE------YDEKFKAFQDDLQKWEET 116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       YSAEFALTENSLKAQAQAQYQMNRPQLIPASVQMPIGINTA
YHQKSLTPSETSASATPPAHRPRAATTVG -> EMLGTLLN
AKYQFPQFPQAIQQNPLLMNNAQHQLILQQQQLAMQHAQAQA
AGSAGTSSSSSSTGTGTSSSRS (in isoform a).
/FTId=VSP_003306.
/FTId=VSP_003307.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            380 SEQENLEKOMR-QLAVIPPMLYDADQQRIKFINMNGLMADPMKYYKDRQVMMM---WSEQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EKETFREKFMQHPKNFGLIASFLERKTVAECVLYYYYLTKKNENYKSLVRRSYRRRGKSQQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QQQQQQQQQQQQPMPRSSQEEK------BEKEKEKEKEKEREK-----EE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRRRAKESKVREYYEKQFPEIRKQRELQERMQSRVGQRGSGLSMSAARSEHEVSEIIDGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels 779; Gaps
                                                ပ
                                                oŧ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 1780;
          Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
Wohldman P.;
"2.2 Mb of contiguous nucleotide sequence from chromosome III
R., Watson A., Weinstock L., Wilkinson-Sproat J.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1780 AA; 198569 MW; 0364817C41A8AD3F CRC64;
                                                                                                                                                                                 Waterston R.;
Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases
-!- SUBCELLULAR LOCATION: Nuclear (Probable).
-!- ALTERNATIVE PRODUCTS:
                                                                                                                                                                                                                                                                                                                                                                            003307;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MARTY SMOOTT; SANTY 2.
PROSITE; PS00034; WYB 1; FALSE NEG.
PROSITE; PS00034; WYB 2; FALSE NEG.
PROSITE; PS50090; WYB 3; 1.
Hypothetical protein; DNA-binding; Nuclear protein;
                                                                                                                                                                                                                                                                                                                                                                                              Note=No experimental confirmation available; SIMILARITY: Contains 1 Myb-like domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
3.3%; Score 438.5; DB 1;
Best Local Similarity 19.1%; Pred. No. 2.4e-07;
Matches 423; Conservative 281; Mismatches 737;
                                                                                                                                                                                                                                                                          Event=Alternative splicing; Named isoforms=2;
                                                                                                                                                                                                                                                                                                                                                                            VSP
                                                                                                                                                                                                                                                                                                                                                                            IsoId=P34333-2; Sequence=VSP_003306,
                                                                                                                                                                                                                                                                                                                              IsoId=P34333-1; Sequence=Displayed;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PIR; S44758; S44758.
WormPep; C14B9.6a; CE00077.
WormPep; C14B9.6b; CE29577.
Interpro; IPR001005; Myb_DNA_binding.
Pfam; PF00249; myb_DNA-binding; 2.
SMART; SM00717; SANT; 2.
                                                                                                                                        [2] REVISIONS, AND ALTERNATIVE SPLICING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; L15188; AAA27949.3; -. EMBL; L15188; AAA27950.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1780
                                                                                                      Nature 368:32-38(1994)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Alternative splicing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1019
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DNA BIND
VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                       elegans.
```

: | | | | KQAQYP---ALQASPHQAA 1184 APRVARPERPRADTGHAFL 1960 PELGKPROSPLTYEDHGAP 1549 rstp--relaksphstvpe 1602 DAAAAYYLPRHLAPNPTYP 1662 AQRADMLRGLSPRESSLAL 1722 LOQQQQHHGTQPEQKS--- 1295 APQPFSSR-HSSSPLSPGG 1781 | | | | | | : | | --GPSSSRIQSGRGVSPA- 1334 SMKGIITAVEPSKPTVLRS 1900 SKP-----FSIQ 2034 | | QONPQQYANLSAAEKLALQ 1563 --LPESQPSSSPLLQTAPG 2125 SVLGGGEDGIEPV---SPP 2228
:: :||:|: ||
PLVVQAQNGIQPIKDYGPP 1767 SATPQVKRIALSQSSPVQR 1234 HAPIWRPGTEQSSGSS 1841 ----VSGHATIARTPAKN 2001 VIKLISSSEAAAAANPSAK 1503 SLDKSHLEGELRPK----Q 2090 ---KFHL---LRPNAEIVR 1611 A----PLPAPLYSFPGASC 2171 czonin) (Brain-; Euteleostomi; e; Murinae; Mus. VE SPLICING,

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           "Isold=O9QYX7-2; Sequence=VSP_003928, VSP_003929;
-1- TISSUE SPECIFICITY: Highly expressed in brain. Low levels found in stomach. Not detected in other tissues analyzed including adrenal gland, testis and pancreas.
-1- DOWAIN: C2 domain 1 is involved in binding calcium and phospholipids. Calcium binds with low affinity but with high specificity and induces a large conformational change.
-1- SIMILARITY: Contains 1 PDZ/DHR domain.
                                                                      "Aczonin, a 550-kd putative scaffolding protein of presynaptic active zones, shares homology regions with rim and bassoon and binds profilin.";
                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE=22384373; Pubbed=12401783; Full Matsumoto M., Fujimoto K., Shibasaki T., Yokoi N., Kashima Y., Matsumoto M., Fujimoto K., Tajima N., Iwanga T., Seino S., Sasaki T., Tajima N., Iwanga T., Seino S., Seino S., Sasaki T., Tajima N., Iwanga T., Seino S., Cipa C., Ca2+ sensor in pancreatic beta-cells. Involvement of cAMP-GEFII.Rim2.Piccolo complex in CAMP-dependent exocytosis."; J. Biol. Chem. 277:50497-50502 (2002).

J. Biol. Chem. 277:50497-50502 (2002).

-i. FUNCTION: May act as a scaffolding protein involved in the organization of synaptic active zones and in synaptic vesicle
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             trafficking.
SUBUNIT: Interacts with Rabaci/Prai, RIMS2 and profilin.
SUBCELLULAR LOCATION: Concentrated at the presynaptic side of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GO:0045202; C:synaptic junction, IDA.
GO:0045202; C:synaptic junction, IDA.
GO:0005544; F:calcium ion binding; ISS.
GO:0005547; F:calcium-dependent phospholipid binding; ISS.
GO:0005543; F:profilin binding; IDA.
GO:000532; F:profilin binding; IDA.
GO:0007019; P:cytcokeleton organization and biogenesis; IDA.
GO:0007013; P:insulin secretion; IDA.
GO:0017157; P:regulation of exocytosis; IDA.
GO:0017157; P:regulation of exocytosis; IDA.
GO:0017040; P:synaptic vesicle targeting; NAS.
                                                   Wang X., Kibschull M., Laue M.M., Lichte B., Petrasch-Parwez E.,
Kilimann M.W.;
                                                                                                                                                                                                                                                                                                            TISSUE=Brain;
Huang W., Jin W., Huang C., Chen B., Zhang J., Ju G.;
Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                  Kilimann M.W.;
Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Event=Alternative splicing; Named isoforms=2;
SPECIFICITY, AND INTERACTION WITH PROFILIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Name=1;
IsoId=Q9QYX7-1; Sequence=Displayed;
                           TISSUE=Brain;
MEDLINE=99439764; PubMed=10508862;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR000008; C2.
InterPro; IPR001478; PDZ.
InterPro; IPR008899; Znf_piccolo.
Pfam; PF00168; C2; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL, Y19185; CAB60731.2; -.
EMBL, Y19186; CAB60732.2; -.
EMBL, AF181269; AAD55786.2; -.
HSSP, P04410; 1A25.
                                                                                                                                                                  Cell Biol. 147:151-162(1999).
                                                                                                                                                                                                                                                                                       SEQUENCE OF 4502-4682 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      synaptic junctions.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           :1349390; Pclo.
                                                                                                                                                                                                                                                                                                                                                                                           INTERACTION WITH RIMS2
                                                                                                                                                                                                               REVISIONS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MGD; MGI
        TISSUE
```

```
1259 REDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETAAPKRTYDMMEGRV 1318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            KKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEASVLRGTALGSVPGGSITK----- 1210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ------GIPSTRVPSDSALTYRGSITHGTPADVLYKGTITRIIGEDSPSRLDRG 1258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1019 -----PRGKSRSPAPPADKEAFAAE-----AQKLPG--DPPCWTSGLPFPVP----- 1058
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ----PREVIKASPHAPDPSAFSYAP-----PGHPLFLGLHDTARPVLPRP--PT1S 1103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    965 -QLKQRAAAIPPIQVTKVHEPPREDAAPTKPAPPAPPPPQNLQPESDAPQQPGSS---- 1018
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           392 ТКРЅРООРГРАКРОРООРVATКРОРООРАРАКРОРОНРТРА----КРОРООРТРАКРОРО 447
                                                                                                                                                                                                                                                                                                                                              91;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       227 -IPQDIPSKSVSS-----QQAEKTKPQAPGTAKPSQQSPAQTPAQQAKPVAQQPGPAK 278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               219 ATVQQPGPAKSPAQ-----PAGTGKSPAQPPVTAKPPAQQAGLEKTSLQQPGPKSLAQT 332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               333 PGQGKVPPGPAKSPAQQPGTAKLPAQQPGPQTASKVPGPTKTPAQLSG-PGKTPAQQPGP 391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAE-----KKEGGSGRATTAKSS 909
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              910 GAPODSDSSATCSADEVDEAEGGDKNRLLSP---RPSILTPTGDPRANASP--OKPLDLK 964
                                                                                                                                                                                                                                                                                                                                                                                                                                               1104 NPPPLISSAKHPSV----LERQIGAISQGMSVQ-LHVPYSEHAKAPVGPVTWGLPLPMDP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        499 KTICPLCNTTELLLH-----TPEKANFNTCTECOSTVCSLCGFNPNPHLTEIKEWLCLN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             553 COMORALGGELAAİPSSPOPTPKAASVOPATASKSPV--------PSQQASP
                                                                                                                                                                                                                                                                                                                                                                                                                               748 TESIPSPHTEAAKDIGQNGPKPPATLGADGPPPGPPTPPRR-TSRAPIEPTPASEATGAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    807 TPPPAPPSPSAPPPVVPKEEKEEETAAAPPVEEG--EEQKPPAAEELAV-----DTGKA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             122 LÞGRSPSTISLKESKSRTDFKEEYKSSMMP---ĠFFSEVNÞLSAVSSVVNKFNPFDLISD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            688 AAASEEAAFPPVVEDEEMEASGVSGNEEEMVEEAEALHASGNEVPRGECSGPATVNNSSD
                                                                                                                                                                                                                                                                                                                                                                                           10 X 10 AA TANDEM APPROXIMATE REPEATS OF
                                                                                                                                                                                                                                                                                                                                              Indels 663; Gaps
Pfam; PF00595; PDZ; 1.
Pfam; PF05715; Zf_piccolo; 2.
SMART; SM00239; CZ; 2.
SMART; SM00228; PDZ; 1.
PROSITE; PS00499; CZ_DOMAIN 1; 1.
PROSITE; PS00404; CZ_DOMAIN 2; 2.
PROSITE; PS50106; PDZ; 1.
Calcium/phospholipid-binding; Metal-binding; Zinc; Zinc-finger;
                                                                                                                                                                                                                                                                                                 3.3%; Score 434.5; DB 1; Length 5038; Local Similarity 19.8%; Pred. No. 8.6e-07; es 391; Conservative 214.
                                                                                                                                                                                          PDZ.
C2 DOMAIN 1.
C2 DOMAIN 2.
TXPTN - SKRRK (in isoform 2).
/FTIG=VSP 003928.
Missing (in isoform 2).
/FTIG=VSP 003928.
                                                                                                                                                                                                                                                                                     5038 AA; 547600 MW; DADA460CF3B40888 CRC64;
                                                                                                                                       P-A-K-P-Q-P-Q-P-X.
C4-TYPE (POTENTIAL).
C4-TYPE (POTENTIAL).
                                                                                                                                                                                 POLY-PRO.
                                                                                                            Repeat, Alternative splicing
DOMAIN 371 470
                                                                                                                                                        526
990
2329
4488
4705
5012
                                                                                                                                                                                                                                                                  5038
                                                                                                                                                        502
967
2305
4394
4607
4922
4829
                                                                                                                                                                                                                                                                     4834
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1059
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  859
                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                         VARSPLIC
                                                                                                                                                                                                                                                                     VARSPLIC
                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                        ZN FING
                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                                                                                            Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
    à
```

us-09-522-753-5.rsp

අ	S97 KKE-LP	ė	
ò	1319 GRAISSASIEGLMGRAIPPERHSPHHLKEQHHIRGSITQGIPRSYVEAQEDYLRREAKLL 1378	ें ह	
QQ	609	g 2	1
ò	1379 KREGTPPPPPPSRDLTEAYKTQALGPL-KLKPAHEGLVATVKEAGRSIHEIPREELRHTP 1437	Š 1	
ΩP	638 -AEALPKPAPPKKP-SAALPEQAKAPVADVEPKQPKTTETLTDSPSSAAATSKPAI 691	<u>a</u> 8	
ò	1438 ELPLAPRPLKEGSITQGTPLKYDTGASTTGSKKHDVRSLIGSPGRTFPPVHPLDVMADAR 1497	Š 1	
qq	692LSSQVQAQAQVTTAPPLKTDSAKTSQSPPPTGDTITPLDSK 732	Q C	1494 TESOKRE
ò	1498 ALERACYBESLKSRPGTASSSGGSIARGAPVIVPELGKPRQSPLTYEDHGAPFAGHLPRG 1557	RESULT	MT 19
đ	PKPDTKPVP	d o s	SHK1 RAT
ò	1558 SPVTWREPTPRLQEGSLSSSKASQDRKLTSTPREIAKSPHSTVPEHHPHPISPY 1611	E I	28-FEB-2003 (Rel
qa		666	28-FEB-2003 (Rel 28-FEB-2003 (Rel
ò	1612 EHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIR 1671	306	interacting prot
ga	833 ETVTGKLFG	3 8	SHANKI.
ò	1672 GYPDTAALENRQTIINDYITSQQMHHNTATAMAQRADMLRGLSPRESSLALNYAAGPRGI 1731	383	Kattus norvegicu Eukaryota; Metaz
οg	877	88	Mammalla; Euther NCBI_TaxID=10116
ζ	1732 IDLSQVPHLPVLVPPTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTHLTKPTTT 1791	R R R	SEQUENCE FROM N.
qq	888PKSTG	χ Σ	DLG4. TISSUE=Brain;
ò	1792 SSSERERDRDRERDREREKSILJSTTTVEHAPIWRPGTEQSSGSSGSSGGGGSSSRP 1851	χ <u>γ</u>	MEDLINE=99419021 Yao I., Hata Y.,
q	::	RT	"Synamon, a nove protein 90/posts
ò	1852 ASHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGIITAVEPSKPTVLRSTSTSSPVRPAA 1911	R.L.	J. Biol. Chem. 2 [2]
q		RP C	SEQUENCE FROM N.
i è	TPPDPATHCP-ICCTIONCYPTIMEDVII.DRRADBURDBURDBURDBURDBURDBURDBURDBURDBURDBUR	. Z 6	MEDLINE=99360650
3 6	NEWATCH CONTRACT CONTRACT OF THE CONTRACT OF T	2 2 5	Weinberg R.J., W
3 6	או זו כי בי הכינת לא כיות כי בי בי ווח בי בי החים היא להיים בי בי בי בי מיוב בי	RT	the NMDA recepto
õ		r r	Neuron 23:569-58 [3]
đ	1037 SSGPKASPVPAPAEPPPQKTPTAAHAKGKKKETEVKAETEKQIPEKETPSIEKTPPAV 1094	ጸ የ	SEQUENCE FROM N.
ò	DPHREKTQSKPFSIQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTH 	Z S	MEDLINE=20549637
qa	1095 ATDQKLEESEVTKSLVSVLPEKKPS-EEEKALPADKKE- 1131	5 E I	"The G protein-c
È	2078 DKSHLEGELRPKQPGPVKLGGEAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQ 2130	7 Z Z	J. Biol. Chem. 2
qq	1132KKPPAAEAPPLEEKKPIPDDQKLPPDAKPSASEGEEKDLLKAHV 1176	X X I	PARTIAL SEQUENCE
ò	2131 RVVTLACHISEVITQDYTRHHPQQLSAPLPAPLYSFPGASCPVLDLRRPPSD 2182	R R	DEVELOPMENTAL SI TISSUE=Brain;
QQ	::	1 & K	MEDLINE=99436166 Lim S., Naisbitt
ò	2183 LYLPPPDHGA-PARGSPHSEGGKRSPEPNKTSVLGGGEDGIEPVSPPEGMTEPGHSRSAV 2241	RT	"Characterizatio genes, alternati
අු	: : : : : : : : : : : : : : :	RT RL	development."; J. Biol. Chem. 2
ò	2242YPLLYRDGE 2250	Ř Ř	[5] PARTIAL SEQUENCE
g	1277 SGSLGEIPSLIPSDEKDLIKGLKKDSFSQESSPSSPSPLAKLESTVLSILEAQASTLVGE 1336	Z X	TISSUE=Brain; MEDLINE=20020275
ò	2251QTEPSRMGSKSPGNTSQPPAFFSKLTESNSAMVKSKKQEINKKLNT 2296	X X	Zitzer H., Hoenc "Somatostatin re
qq	: : :: :: : : : : : : : : : : : : : : : : : : : : : : : : : : : :	RT	multidomain prot J. Biol. Chem. 2

```
2342
                                                                                                               2401
                                                      PELVDDESPRRASYDSVEDSSESENSPVARRKRRTSIGSSSSEEYKQEDSQGSGE 1456
                                                                                                                                             STANDARD; PRT; 2167 AA.

1. O9WUB8;
1. 141, Created)
1. 1. Last sequence update)
1. 141, Last annotation update)
1. 141, Last annotation sprotein 1 (Shank1) (GKAP/SAPAP otein) (SPANK1) (Synamon) (Somatostatin receptor otein) (SSTR interacting protein) (SSTRIP).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           21; PubMed=10488079;
., Hirao K., Deguchi M., Ide N., Takeuchi M., Takai Y.;
.el neuronal protein interacting with synapse-associated
ssynaptic density-95-associated protein.";
274:27463-27466 [1999].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       I.A. (ISOFORMS 1 AND 3), AND INTERACTION WITH DLGAP1 AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N.A. (ISOFORM 4), AND INTERACTION WITH DLGAPI.

-Dawley;

50; PubMed=10433268;

im E., Tu J.C., Xiao B., Sala C., Valtschanoff J.,

Worley P.F., Sheng M.;

1 family of postsynaptic density proteins that binds to

tor/PSD-95/GKAP complex and cortactin.";
-----YNISQPGTEIFNMPAI----TGTGLMTYRSQAVQEHASTNMGL
                                                                                                               KALMGKYDQWEESPPLSANAFNPLNASASLPAAMPITAADGRSDHTLTSPGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       66; PubMed=10506216;
it S., Yoon J., Hwang J.I., Suh P.G., Sheng M., Kim B.;
ion of the Shank family of synaptic proteins. Multiple
ive splicing, and differential expression in brain and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             us (Rat).
120a; Chordata; Craniata; Vertebrata; Buteleostomi;
121a; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                37; PubMed=10958799;
edhof T.C., Stahl B.;
coupled receptor CL1 interacts directly with proteins
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    15; PubMed=10551867;

ck H.-H., Baechner D., Richter D., Kreienkamp H.-J.;

receptor interacting protein defines a novel family of

receins present in human and rodent brain.";

274:32997-33001(1999).
                                                                                                                                                                                                                              SRPSSRKAKSPAPGLASGDRPPSVSSVHSEGDCNRRTPLTNRVWED 2452
                                                                                                                                                                                                                                                                  TEROM N.A. (ISOFORMS 1; 2; 3; 4 AND 5), AND STAGE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              E FROM N.A. (ISOFORM 4).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             275;36204-36210(2000)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           274:29510-29518(1999)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        .A. (ISOFORM 1).
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -
                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               . 1174
1246
1323
1331
1726
AA; 226333 M
                                                                                                                                                                                InterPro; IPR002110; ANK.
InterPro; IPR001478; PDZ
InterPro; IPR001660; SAM.
InterPro; IPR001452; SH3.
                                                                                                                                                                                                                                                                                            PD000066; SH3; 1
                                                                                                                                                                                                                                                                                                                                                                                                        SH3;
SAM D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1943
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     932
1015
1027
1199
1860
613
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2167
                                                                                                                                                                                                                                                   Pfam; PF00595; PDZ; 1.
Pfam; PF00536; SAM; 1.
Pfam; PF00018; SH3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          654
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    804
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    654
                                                                                                                                                                                                                                         Pfam; PF00023; ank;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity
                                                                                                                                                                     P00519; 1AB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1930
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1246
1323
1331
1726
2167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           646
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1944
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    797
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1141
                                                                                                                                                                                                                                                                                                                                                                                                                                     ANK repeat;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                         PROSITE;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                            PROSITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                               REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
    ISOId=O9WV48-5; Sequence=VSP 006076, VSP 006077;
TISSUE SPECIFICITY: Expressed only in brain (neuropil of cortex, CA1 region hippocampus and molecular layer of cerebellum be bevelopWENTAL STAGE: Expression increases from low levels at birth to high levels at 3-4 weeks before dropping slightly in adulthood. Expressed in the cortex and the molecular layer of the cerebellum at postnatal day 7. Isoform 2 expression does not change during developmment of both cortex and cerebellum. Isoform 4 expression decreases significantly during development of cortex but not
                                                                                                                                                                                                                                                                                                                                                                             Sheng M., Kim E.;
"The Shank family of scaffold proteins.";
"The Shank family of scaffold proteins.";
J. Cell Sci. 113:1851-1856 (2000).

-1- FUNCTION: Seems to be an adapter protein in the postsynaptic density (PSD) of excitatory synapses that interconnects receptors of the postsynaptic membrane including NMON-type and metabotropic glutamate receptors, and the actin-based cytoskeleton. May play a role in the structural and functional organization of the dendritic spine and synaptic junction. Overexpression promotes maturation of dendritic spines and the enlargement of spine heads via its ability to recruit Homer to postsynaptic sites, and enhances presynaptic function.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SUBUNIT: May homomultimerize via its SAM domain. Interacts with SPTANI, Homer-I and DLGAPI/GKAP. Is part of a complex with DLG4/PSD-95 and DLGAPI/GKAP. Interacts with SSTR2 C-terminus via
              INTERACTION WITH HOWER-1, AND SUBCELLULAR LOCATION.
MEDLINE=99360651; PubMed=10433269;
Tu J.C., Xiao B., Naisbitt S., Yuan J.P., Petralia R.S., Brakeman P.,
Doan A., Aakalu V.K., Lanahan A.A., Sheng M., Worley P.F.;
"Coupling of mGluK/Homer and PSD-95 complexes by the Shank family of
Dostsynaptic density proteins.";
Neuron 23:583-592(1999).
                                                                                                                                                                                     the
                                                                                                                                                                                                                                                                     MEDLINE=21389514; PubMed=11498055;
Sala C., Piech V., Wilson N.R., Passafaro M., Liu G., Sheng M.;
"Regulation of dendritic spine morphology and synaptic function by
Shank and Homer.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     the PDZ domain (By similarity).
SUBCELLULAR LOCATION: Cytoplasmic; postsynaptic density of neuronal cells. Colocalizes with alpha-latrotoxin receptor 1.
                                                                                                                                         MEDLINE=21523912; PubMed=11509555;
Bockers T.M., Mameza M.G., Kreutz M.R., Bockmann J., Weise C.,
Buck F., Richter D., Gundelfinger E.D., Kreienkamp H.-J.;
"Synaptic scafffolding proteins in rat brain. Ankyrin repeats of
multidomain Shank protein family interact with the cytoskeletal
protein alpha-fodrin.";
J. Biol. Chem. 276:40104-40112(2001).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           cerebellum.
SIMILARITY: BELONGS TO THE SHANK FAMILY.
SIMILARITY: Contains 7 ANK repeats.
SIMILARITY: Contains 1 PDZ/DHR domain.
SIMILARITY: Contains 1 sterile alpha motif (SAM) domain.
SIMILARITY: Contains 1 SH3 domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 IsoId=Q9WV48-2; Sequence=VSP_006072, VSP_006073;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Event=Alternative splicing; Named isoforms=5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Comment=Additional isoforms seem to exist;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IsoId=Q9WV48-3; Sequence=VSP_006074;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        IsoId=09WV48-4; Sequence=VSP_006075;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IsoId=Q9WV48-1; Sequence=Displayed;
                                                                                                                                                                                                                                                                                                                                                                    MEDLINE=20267867; PubMed=10806096;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ALTERNATIVE PRODUCTS:
                                                                                                                                                                                                                                                                                                                            Neuron 31:115-130(2001)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Name=4; Synonyms=A;
                                                                                                                                INTERACTION WITH SPIANI
          [6]
INTERACTION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Name=1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Name=3
                                                                                                                                                                                                                                                         FUNCTION
                                                                                                                                                                                                                                                                                                                                                        REVIEW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       + +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -
```

```
87;
use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SQEGRQESRSDKAKRLFRHYTVGSYDSFDAPSLIDGIDSG
-> MALSAVGGGGGGALPQPPPALSSSWPALGPRRRSVWY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Missing (In isoform 4).
/FIId=VSP 006075.
LSEDSQTSLLSKPS -> QYRIVVKSSDFGDF (in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3.3%; Score 432; DB 1; Length 2167;
21.0%; Pred. No. 4.7e-07;
ive 139; Mismatches 619; Indels 722;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 2167;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A -> T (IN REF. 1).
S -> D (IN REF. 1).
S -> N (IN REF. 2).
MW; 3F478B5A7B18BA86 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Alternative splicing.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Missing (in isoform 3). /FIId=VSP_006074.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (in isoform 2)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Missing (in isoform 5) /FIId=VSP_006077.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Y (in isoform 2) FTId=VSP 006073.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             006076.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               006072
                                                                                                                                                                                                                                                                                                                                                                                                        Prodom; ruccossistant SMO248; ANK; 6.
SMART; SMO0248; PDZ; 1.
SMART; SMO0454; SAM; 1.
SMART; SMO0326; SAN; 1.
PROSITE; PSS0088; ANK REPEAT; 3.
PDOSITE; PSS0089; ANK REPEAT; 3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (soform 5)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FTIG=VSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /FTIG=VSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ANK 1.

ANK 2.

ANK 3.

ANK 4.

ANK 5.

ANK 6.

ANK 7.

SH3.

PDZ.

PDZ.

PDZ.

POLY-PRO.

POLY-HIS.

POLY-HIS.

POLY-GLY.
                                                                                                                                               EMBL; AF102855; AAD04569.2; -.
EMBL; AF131951; AAD23417.1, ALT_INIT.
EMBL; AF159046; AAD42975.1; -.
EMBL; AF141904; AAF02498.1; ALT_INIT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative 139;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Sim
Matches 394;
```

Qy 1538 QSPLTYEDH :	Qy 1718 SSIALNYAAQ Db 1635 SEVATLTQGA Qy 1778 SPGGPTHLTI Db 1670GPDI Qy 1838 SGSSGGGGGG Qy 1838 SGSSGGGGGG Qy 1898 LR6 Qy 1898 LR6	1951 1951 1852 2066 1911 2118 1951 2073 2015 2045	RESULT 20 MLL4 HUMAN ID MLL4 HUMAN ID MLL4 HUMAN STAN AC Q9UNN6; 015022; 099 DT 16-OCT-2001 (Rel., DT 10-OCT-2001 (Rel., DT 10-OCT-2003 (Rel., DE Myeloid/lymphoid on DE Myeloid/lymphoid on DE Myeloid/lymphoid on DE MLM OR TRX2 OR HRS OS HOMO Sapiens (Human OC EUKaryota; Metazoa, OC Mammalia; Eutheria OX NCBI_TAXID=9606; RN [1] RP SEQUENCE FROM N.A.
	PKPPATLGADGPPGPPTPPRRTSRAPIEPTPASEATGAPTPPPAPPSPSA	988 ĠĠĠŖĔĸŚĿ	GSSSAFTSFLPRPLVHPLTGKALDPASPLGLALAARERALKESSEGGGTPQPPPREP
6 6 6 6 6	8 6 8 6 8 6 8	8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	6 6 6 6 6 6 6 6

1597	1657 1611	1717	1777	1837	1714	1897 1757	1950 1796	2010	2065	1910	2117	2172	2000	2214	2044	26	202			
8 QSPLTYEDHGAPFAGHLPRGSPVTWREPTPRLQEGSLSSSKASQDRKLTSTPREIAKSPH 	8 STUPEHHPHPISPYEHLLRGVSGVDLYRSHIPLAFDFTSIPRGIPLDAAAAYYLPRHLAP :	8 NPTYPHLYPPYLIRGYPDTAALENRQTIINDYITSQQMHNNTATAMAQRADMLRGLSPRE	B SSLALNYAAGPRGIIDLSQV	S SEVALLI COMEMNEGL	:	8 SGSSGGGGGSSSRPASHSHAHQHSPISPRTQDALQQRPSVLHNTGWKGIITAVEPSKPTV	B LRSTSTSSPVRPAATFPPATHCPLGGTLDGVYPTLMEPVLLPKEAPRVARPER	1. PRADTGHAFLAKPPARSGLEPASSPSKGSEPRPLVPPVSGHATIARTPAKNLAPHHASPD	1 PPAPPASASDPHREKTQS-KPFSIQE	.2 PPPPPLPGPLSQPQASALATVKASIISELSSKLQQFGGSS-TAGGALPWARGGSGGSTDS	6 HDKGLPKHLEELDKSHLEGELRPKQPGPVKLGGEAAHLPHLRPLPESQPSSS	8PLLQTAPGVKGHQRVVTLAQHISEVITQDYTRHHPQQLSAPLPAPLYSFPGASCP	.1 PKPPLPPLPTGSGVSSSTAAAPGATSPSASSASASTR-HLQGVEFEMRPPL	3 VLDLRRPPSDLYLPPPDHGAPARGSPHSEGGKRSPEPNKTSV	1LRRAPSPSLLPASDH	5LQGGEDGIEPV-SPPEGMTEPCHSRSAVYPLLYRDGEQT	S PTGGGGGSTDPFAPVFVFPHFGISGGLGGALSGASKSLSFTKLLSLF	3 GNTSQPPAFFSK 227	3 DKP	LT 20 HUMAN MIL4 HUMAN STANDARD; PRT; 2715 AA. MIL4 HUMAN STANDARD; PRT; 2715 AA. MIL4 HUMAN STANDARD; PRT; 2715 AA. 99UMN6; 0015022; 095836; 0966P2; 0961P3; 09UK25; 09Y668; 09Y669; 16-OCT-2001 (Rel. 40, Last sequence update) 10-OCT-2003 (Rel. 42, Last annotation update) 10-OCT-2003 (Rel. 42, Last annotation update) Myeloid/lymphoid or mixed-lineage leukemia protein 4 (Trithorax homolog 2). Myeloid/lymphoid or MiL2 OR KIAA0304. Homo sapiens (Human). Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. NCBI_TAXID=9606; [1] SEQUENCE FROM N.A. (ISOFORM 1).
153	159	165	171	177	167	183	189	195	0	185	206	211	195	217	200	2 0	9	0 0	209	HUMAN MIL4 HUMAN MIL4 HUMAN O9UWN6; 0150 16-0CT-2001 16-0CT-2001 10-0CT-2009 Myeloid/Lymp homolog 2) Mil4 OR TRXZ HOMO sapiens Eukaryota; M Mammalia; Eu NCBI_TaxID=9 [1] SEQUENCE FRC
දු දු	දු දු	중 음	કે ક	3 &	qq	දි සි	දු දු	දි දි	ò	q	중 원	δ	qq	ò	ДG	8 1	g ·	ð 1	g	RESULALION DE

```
Also found in brain, bone marrow, heart, muscle, kidney, pancreas,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        365
1005
1252
1303
1396
1471
2695
2715
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1201
1249
1335
1449
2574
2699
2699
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; AB002302;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AD000671;
                                                                                                                                                                                                                                                                                              AC 014686).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
    TISSUBBRAIN, and Skin;

**REQUENCE OF 1918-2715 FROM N.A.*

**REQUENCE-23825.*

**REQUENCE-23825.*

**REAURINE-23825.*

**RETAINDEDET R.D., Feingold E.A., Grouse L.H., Derge J.G.,

**RIACHUL S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,

**RIACHUL S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,

**RIACHUL S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,

**POSTAGE M.B., Bonaldo M.F., Rang C.P., Haibe F.,

**Rapleton M., Soares M.B., Bonaldo M.F., Carninci P., Prange C.,

**Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

**Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

**Villaion D.K., Wuzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

**Postage S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

**Villaion D.K., Wuzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

**Postage S., Worley W.C., Shevchenko Y., Bouffard G.G.,

**Phitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

**Rachiguez A.C., Grimwood J., Schmutz J., Myers R.M.,

**Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

**Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,

**Roberztion and initial analysis of more than 15,000 full-length
                                                                                                                                                                                                                                                                                                                                                                                   TISSUE-Leukocyte, and Testis;
MEDLINE=20105772; PubMed=10637508;
Huntsman D.G., Chin S.-F., Muleris M., Batley S.J., Collins V.P.,
Wiedemann L.M., Aparitio S., Caldas C.;
"MLL2, the second human homolog of the Drosophila trithorax gene, maps
to 19q13.1 and is amplified in solid tumor cell lines.";
Oncogene 18:7975-7984(1999).
                                                                                                                                                                                          Lamerdin J.E., McCready P.M., Adamson A.W., Burkhart-Schultz K., Garcia E., Kyle A., Ramirez M., Stilwagen S., Garnes J., Danganan L., Bluce R., Quan G., Montgomery M., Ow D., Kobayashi A., Olsen A.O., Carrano A.V.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE=97349984; PubMed=9205841;
Nagasea T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,
Nagasea T., Tankawa A., Foctani H., Nomura N., Ohara O.;
"Prediction of the coding sequences of unidentified human genes. VII.
The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro.";
DNA Res. 4:141-150(1997)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Name=2; Synonyms=Truncated;
Isold=Q9UMN6-2; Sequence=VSP 006668, VSP 006669;
TISSUE SPECIFICITY: Widely expressed. Highest levels in testis.

    -1- FUNCTION: Possibly acts as a transcriptional regulatory factor.
    -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
    -1- ALTERNATIVE PRODUCTS:

Angrand P.O., Valvatne H., Jeanmougin F., Adamson A., van der Hoeven F., Olsen L., Tekotte H., Huang N., Poch O., Lamerdin J., Chambon P., Losson R., Stewart A., Aasland R.; "Mammalian trithorax- and ASH1-like proteins: putative chromatin regulators which contain PHD fingers and SET domains."; Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    "MLL2: A new mammalian member of the trx/MLL family of genes.", Genomics 59:187-192(1999).
                                                                                                                                                                                                                                                                                           "Sequence analysis of a 1 Mb region in human 19q13.1."; Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     human and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Event=Alternative splicing; Named isoforms=2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PARTIAL SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
TISSUE-Bone marrow, and Placenta;
MEDGLINE-99339983; PubMed=10409430;
FitzGerald K.T., Diaz M.O.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE OF 816-2715 FROM N.A. (ISOFORM 1).
                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE OF 111-2715 FROM N.A. (ISOFORM 1)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Name=1; Synonyms=Long;
IsoId=Q9UMN6-1; Sequence=Displayed;
                                                                                                                                                                         SEQUENCE FROM N.A. (ISOFORM 1)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TISSUE=Brain;
    REAL TARKET TARK
```

<u>-</u>

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its most by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GO; GO:0005634; C:nucleus; NAS.
GO; GO:0003700; F:transcription factor activity; NAS.
GO; GO:0003700; F:transcription factor activity; NAS.
GO; GO:00048096; F:rinc ion binding; NAS.
GO; GO:0048096; P:chromatin-mediated maintenance of transcrip. . .; NAS.
InterPro; IPR003889; FYrich. C.
InterPro; IPR00388; FYrich. N.
InterPro; IPR00366; PostSST.
POST-SET.
POLY-GLV.
POLY-PRO.
ASP/GLU-RICH (ACIDIC).
PRO-RICH.
POLY-GLN.
POLY-PRO.
POLY-PRO.
                                                                                                                                                                                                                                                                                                                                                                    -; NOT ANNOTATED CDS.
BAA20763.2; -.
                                                                                                                                                                                                                                                                                                                                                                                                     EMBL, AF186605, AAD56420.1;
EMBL, AF104918; AAD1732.1;
EMBL, AF106279; AAD26113.1;
EMBL, BC009337; AAH09337.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; BC007353; AAH07353.1; -. EMBL; AF105280; AAD26112.1; -. MIM; 606834; -.
                                                                                                                                                                                                                                                                                                                                                      EMBL; AJ007041; CAB45385.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
```

ą	::: 788 VEEKMFS	 VEEKMFSLLKRAKVQLFKIDQQQQKVAASMPLSPGGQMEEVAGAVKQISDRGPVRSEDE 847
À	1178	AGPPESLGVPTAQEASVLRGTALGSVPGGSITKGIPSTRVPSDSAIT 1224
ą	848 SVEAKRE	
λ	1225 YRGSITH	YRGSITHGTPADVLYKGTITRIIGEDSPSRLDRGRED1261
ą	898 LADR	QDLATEDTSSASETESVPSRSRRĞKVEAAGPGGESEPTGSGGTLAHTPRR 951
દ્ર ક	1262 SLPKGHVIYEGKK	IYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGP-PHETAAPKRT 1310
æ	1311 YDMMEGR	VGRAISSASIEGLMGRAIPPERHSPHHLKEQHHIRGSI 1355
ą	1005 CDKIEAR	CDKIEARKMERLAKKGRTIVKTLLPWDSDESPEASPGPPGPRRGAG 1050
≿	1356 TQGIPRSYV	YVEAQEDYLRREAKLLKR 1380
á	•	AGG-PREEVVAHPGPEEQDSLLQRKSARRCVKQRPSYDIFEDSDDSEPGGPPAPRRTPR 1109
≱ a	1381 EGTPPPPPSR 	PPSR1391 BPROSRPRYPTI OPULOI KARRRI DKDALAPGEPRASPBNIWTGKOKSPDIGUH 1169
≿	1392	
ą	1170 RVRVDFK	RVRVDFKEDCDLENVWLMGGLSVLTSVPGGPPMVCLLC-ASKGLHELVFCQVCCDFFHPF 1228
⋩	1442APR	APRPLKEGSITQGTPLKYDT16ASTTGSK1469
ą	1229 CLEEAERPI	
<i>\</i>	1470	KHDV1500
ą	1280 LGPSYPT	LGPSYPTRATRKRRHWICSACVRCKSCGATPGKNWDVEWSGDYSLCPRCTQLYEKGNYCP 1339
≿:	1501RACYE	
ą	1340 ICTRCYE	CTRČÝEDNDYESKAMĄCAĄCDHWVHAKCEGLSDEDYETĽSGLPDSVLYTCGPCAGAAQP 1399
≿ :		SPVTMREPTPRLQE 157
g	1400 RWREALS	RWREALSGALQGGLRQVLQGLLSSKVVGPLLLCTQCGPDGKQLHPG-PCGLQAVSQRFED 1458
≿ 9	1572 GSLSSSK 	GSLSSSKASQDRKLTSTPREIAKSPHSTVPEHHPHPISPYEHLLRGV 1618
≿		166
ą	1500 LLESAFG	: : : 1. Llesargwpdahdpkywrrstrlpngvlpnavlp-psldhyyaqwrqqepe 1549
≿	1666P	PPYLIRGYPENR 1682
ą	1550 TPESGQPP	
≿:	1683 QTIINDY	QTIINDXITSQQMHHNTATAMAQRADM1709
ð	1606 WTHVNCA	WTHVNCAIWSAEVFEENDGSLKNVHAAVARGRQMRCELCLKPGATVGCCLSSCLSNFHFM 1665
≿	1710	1715 L.
و	1666 CARASYC	CARASYCIFQDDKKVFCQKHTDLLDGKEIVNPDGFDVLRRVYVDFEGINFKRKFLTGLEP 1725
	1716 RESSLALI 	SLALNYAAGPRGIIDLSQVPHLPV
٠ ج	S PPTP	PGGPTHLTKPTTTSSSERERDRDR 180

OR AIB3 OR RAP250 OR TRBP OR KIAA0181.

```
2437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2489
                                                                                                                                                                                                                                                                                                                                                                                                                         2220
                                                                                                                                                                                     RWASP-PIKTSPQLRVPPTSVVTALTPTSGELAPPGPAPSPPPPEDLGPDFEDMEVVSG 1985
                                                                                                                                                                                                               -----PVSPV 2059
                                                                                                                                                                                                                            2165 SFPGASCPVLDLRRPPSDLYLPPPDHGAP--ARGSPHSEGGKRSP-EPNKTSVLGGGEDG 2221
                                                                                                                                                                                                                                                                                                                                                                            2125 GDGGAG------PREESLPP----APPLANGSQPSQGLTASPADPTRTFAWLPGAPG 2171
                                                                                                                                                                                                                                                                                                                                                                                                                                               SKLTESNSAMVKSKKQEINKKLNTHNRNEPEYNISQPGTEIFNMPAITGTGLMTYRSQAV 2332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2246
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2333 QEHASTNMGLEAIIRKA-----LMGKYDQWEESPPLSA-----NAFNPLNASASLPA 2379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ERDRDREREKSILTSTTTVEHAPIWR--PGTEQSSGSSGSSGGGGGSSSRPASHSHAHQH 1860
                                                                                                                                         LGGVSFGPLPSPGSPSSLTHHIPTVGDPDFP-----APPRRSRRPSPLAPRPPPS 1926
                                                                                                                                                                                                                                                            SSPSL--THDKGLPKHLEELD-----KSHLEGELRPK------QPGPVKLGGEAAHLP 2104
                                                                                                                                                                                                                                                                                                         HLRPLPESQPSSSPLLQTAPGVKGHQRVVTLAQHISEVITQDYTRHHPQQLSAPLPAPLY 2164
                                                                                                                                                                                                                                                                                                                               ------KNLGGP---- 2124
                                                                                                                                                                                                                                                                                                                                                                                                                    VRVLSLGPAPBPKPATSKIILVNKLGQVFVKNAGEGEPVPPP------VKQPPL--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |:| | : | |: | |: | | : | | SAPGLAPSATPGAPR-IEQLDGVDDGTDSEAEAVQQPRGQGTPPSGPGVVRAGVLGAAGD
                                                                                                                                                                                                                                                                                                                                                                                                   2222 IEPVS-----PPEGMTEP---GHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQPPAFF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ------PPTISPTAPTSWTLPPGPLLGVLPVV--------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ------GVVRPAPPPPPLTLVLSSGPASPPRQAIRVKRVSTFSGRSPPAPPY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2380 AMPITAADGR-SDHTLTSPGGGGKAKVSGRPSSRKAKSP-APGLASGDRPPSVSSVHSEG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DCNRRIPLINRVWEDRPSSAGSIPFPYNPLIMRLQ----AGVMAS----PPPPGLPAGSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q14686; Q9NTZ9; Q9UR44; Q9UR86; 28-FEB-2003 (Rel. 41, Created) 28-FEB-2003 (Rel. 41, Last sequence update) 28-FEB-2003 (Rel. 41, Last sequence update) 28-FEB-2003 (Rel. 43, Last annotation update) Nuclear receptor coactivator 6 (Amplified in breast cancer-3 protein) (Cancer-amplified transcriptional coactivator ASC-2) (AsC-2) (Activating signal cointegrator-2) (ASC-2) (Peroxisome proliferator-activated receptor-interacting protein) (PARP) (Nuclear receptor-activating protein, 250 kDa) (Nuclear receptor-activating protein, 250 kDa) (Nuclear receptor-activating protein, Cancertivating protein, Cancertivating protein, Cancertivating Cancertivator RAP250) (NRC RAP250) (Thyroid hormone receptor-binding
                                                                     SPISPRIQDALQQRPSVLHNTGMKGIITAVEPSKPTVLRSTSTSSPVRPAATFPPATHCP
                                                                                                                   LGGTLDGVYPTLMEPVLLPKEAPRVARPERPRADTCHAFLAKPPAR----SGLEPASSPS
                                                                                                                                                                  KGSEPRPL-----VPPVSGHATIARTPAKNLAPHHASPDPPAP-----VPPVSGHATIARTPAKNLAPHHASPDPPAP
                                                                                                                                                                                                               2015 -PASASDPHREKTOSKPFSIQELELRSLGYHGSSYSPEGVE----
                                                ---EDPPLDTDVLVPGAPERHSPIONLDPPLRPDSGSA---
 EYRPWGP---REEPAHLEAAEENQTIVHSPAPSSEPPGG-
                                                                                      | :||:
PPPAPRS------FSGARIKVPNYSPSRR-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2063 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       P-LAGPH 2495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PKRTGPH 2413
                                                                                                                                                                                                                                                                                                                                 2102 RARP-PEDLP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HUMAN
                                                                                                                                                                                        1927
                                                                                                                                                                                                                                                                                                                                                                                                                           2172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2407
                                                                                                                                                                                                                                                                                    2043
                                                                                                                                                                                                                                                                                                            2105
                                                                                                                                                                                                                                                                                                                                                                                                                                                   2273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2247
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2490
                                                                                                                      1921
                                                                                                                                           1877
                                                                                                                                                                   1977
                                                                                                                                                                                                                                                             2060
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2221
                                                1818
                                                                                               1853
                          1803
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  protein)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NC06
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8 6 6 6
                                                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                           유
상
음
                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ठ व
                                                                                                              ठ
                                                                                                                                      g
                                                                                                                                                                  강 음
                                             원
                                                                       à
                                                                                       용
                        ò
```

```
REGURENE FROM N.A.

REGURENE FROM N.F., Bartes M.K., Badguley C.L.,

REGURENE M., Bartow M.K.F., Bartes M.K., Bagguley C.L.,

REGURENE M., Bartow M.K.F., Bartes M.M., Brown A.J.,

REGURENE M., Burtill W.D., Butler A.P., Carder C., Carter N.P.,

REGURENE M., Colley V.E., Collark L.M., Clark S.Y., Clee C.M.,

RA Clegg S., Cobley V.E., Collier R.E., Cornor R.E., Garrier P., Rankan D.V., Haarley J.L., Heath P.D., How, J. Honden D.J., Howdel B., Hunt A.R., Hunt S.E., Jekosch K., Johnson D.,

RA M.P., Kimberley A.M., King A., Knightes A., Laird G.K., Lawlor S.,

RAM Marsh V.L., Martin S.L., McConnachie L.J., McLay K., McMurray A.A.,

RAM Marsh V.L., Martin S.L., McConnachie L.J., McLay K., McMurray A.A.,

RAM Marsh V.L., Parter R., Pater R., Plumb R.W., Ramsay H.,

RAM Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,

RAM Phillimore R.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A., HOMODIMERIZATION, AND INTERACTION WITH CREBBP;
RXRA, ESR1; NR3C1; RARA, VDR AND THRA.
MEDLINE=20325329; PubMed=10866662;
MANAJAI M.A., SamuelB H.H.;
A new family of muclear receptor coregulators that integrates nuclear receptor signaling through CBP.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      nuclear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         J.-A.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               motif-containing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             , v
by
                                                                                                                                                                                              TBP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PHOSPHORYLATION BY PRKDC, AND INTERACTION WITH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PPARA; PPARG; ESR1; ESR2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TISSUE=Bone marrow;
MEDLINE=96281124; PubMed=8724849;
Magase T., Seki N., Ishikawa K.-I., Tanaka A., Nomura N.;
Nagase T. oc the coding sequences of unidentified human genes. 
"Prediction of the coding sequences of 40 new genes (XIAA0161-KIAA0200) deduced lanalysis of cDNA clones from human cell line KG-I.";
                                      Euteleostomi;
                                                                                                                                                                                                                                KXRA, ESRI; RARA AND THRA.

MEDLINES-20036574; PubMed=10567404;
MEDLINES-20036574; PubMed=10567404;
Jueg S.-K., Anzick S., Choi J.-B., Bubendorf L., Guan X.-Y.,
Jung Y.-K., Kallioniemi O.P., Kononen J., Trent J.M., Azorsa D.
Jhun B.-H., Cheong J.H., Lee Y.C., Meltzer P.S., Lee J.W.,
A nuclear factor ASG-2, as a cancer-amplified transcriptional
coactivator essential for ligand-dependent transactivation by n
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Treuter E., Gustafsson, a nuclear receptor
                                                                                                                                                                                              AND INTERACTION WITH CREBBP; NCOA1; GTF2A;
                             Bukaryota, Metazoa, Chordata, Craniata, Vertebrata, Butele
Mammalia, Butheria, Primates, Catarrhini, Hominidae, Homo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDIINE-20283376; PubMed=10823961;
Ko L., Cardona G.R., Chin W W.;
Thyyroid hormone receptor-binding protein, an LXXLL protein, functions as a general coactivator.";
Proc. Natl. Acad. Sci. U.S.A. 97:6212-6217(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A., AND INTERACTION WITH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Caira F., Antonson P., Pelto-Huikko M., 1
"Cloning and characterization of RAP250,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Biol. Chem. 274:34283-34293(1999)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    20:5048-5063 (2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Biol. Chem. 275:5308-5317(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE=20148724; PubMed=10681503;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A., PHOSPH
THR; RAR; EP300 AND CRSP3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DNA Res. 3:17-24(1996).
Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TISSUE=Lymphocytes;
                                                                                                                                                                                                       FROM N.A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  [5]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            receptors in vivo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Mol. Cell. Biol.
                                                                                                                      NCBI TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  coactivator.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TISSUE=Testis;
                                                                                                                                                                                                       SEQUENCE
```

Chin W.W.;

Ser-884 adjacent to the LXXLL motif of coactivator TRBP defines

"Ser-884 adjacent to the LXXLL motif of coactivator TRBP defines

"Ser-884 adjacent to the LXXLL motif of coactivator TRBP defines

"Ser-884 adjacent to the LXXLL motif of coactivator TRBP defines

"I mol. Endocrinol. 16:128-140(2002).

"I FUNCTION: Nuclear receptor coactivator that directly binds nuclear receptors and stimulates the transcriptional activities in a hormone-dependent manner. Involved in the coactivation of and AR2-dependent manner. Involved in the coactivation of and prostanoids (PARs and RKs), thyroid hormone (TRS), vitamin D3 (VDR) and prostanoids (PARs). Thyroid hormone (TRS), vitamin D3 (VDR) and prostanoids (PPARS). Thyroid hormone cactivator. May coactivate or rather than just a nuclear receptor coactivator. May coactivate expression vit a remodeling of chromatin and its committee. TISSUE=Cervical carcinoma;
MEDLINE=22371496; PubMed=12482968;
MEDLINE=2371496; PubMed=12482968;
MEDLINE=2371496; PubMed=12482968;
Kam S., Shn Y.C., Kim D.-H., Kim S.-W., Kang M.-J., Jung D.-J.,
Kwak E., Barlev N.A., Berger S.L., Chow V.T., Roeder R.G.,
Azorga D.O., Meltzer P.S., Suh P.-G., Song E.J., Lee K.-J., Lee Y.C., Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E., Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A., Tracey A., Tromans A.C., Vaudin M., Wall M., Walls J.M., Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A. Wilming L., Wray, P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S. MEDDINE=21423955, PubMed=11443112; Iwasaki T., Chin W.W., Ko L.; "Identification and characterization of RRM-containing coactivator activator (COAA) as TRBP-interacting protein, and its splice variant sequence and comparative analysis of human chromosome 20."; SUBUNIT: Monomer and homodiner. Interacts with RNPC2 (By similarity). Interacts in vitro with the basal transcription factors GTF2A and TBP, suggesting an autonomous transactivation function. Interacts with NCOA1, CRSP3, RBM4, the histone acetyltransferases EP300 and CREBBP, and with the methyltransferases NCOA6IP and HRWTILI/PRMT2. Belongs to the ASC-Z/NCOA6 complex (ASCOM), which contains ASC-Z/NCOA6, the retinoblastoma-binding protein RBO-3/ RBBPS, alpha- and beta-tubulins, the trithorax group proteins MLL2 and MLL3, and MEDLINE=21417756; PubMed=11517327; Zhu Y.-J., Qi C., Cao W.-Q., Yeldandi A.V., Rao M.S., Reddy J.K.; "Cloning and characterization of PIMT, a protein with a methyltransferase domain, which interacts with and enhances nuclear receptor coactivator PRIP function."; "Activating signal cointegrator 2 belongs to a novel steady-state complex that contains a subset of trithorax group proteins."; MOI. Cell. Biol. 23:140-149(2003). MEDLINE=21635582; PubMed=1177344; Ko L., Cardona G.R., Iwasaki T., Bramlett K.S., Burris T.P., MEDLINE=22151129; PubMed=12039952; Qi C., Chang J., Zhu Y., Yeldandi A.V., Rao S.M., Zhu Y.-J. "Identification of protein arginine methyltransferase 2 as MUTAGENESIS OF 883-THR--GLU-894, AND PHOSPHORYLATION Proc. Natl. Acad. Sci. U.S.A. 98:10380-10385(2001) INTERACTION WITH MLL3 AND THE ASCOM COMPLEX. coactivator for estrogen receptor alpha."; J. Biol. Chem. 277:28624-28630(2002). Biol. Chem. 276:33375-33383(2001) as a coactivator modulator (CoAM)." INTERACTION WITH HRMT1L1. INTERACTION WITH NCOAGIP Nature 414:865-871(2001) INTERACTION WITH RBM14. "The DNA 급

-1- SUBCELLULAR LOCATION: Nuclear. -!- TISSUE SPECIFICITY: Ubiquitous. Highly expressed in brain,

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              114 SNNQQLRDLGILSVQIE-GEGAINLALAQNRSQDVRMNGPMGAGNSVRMEA-GFPMASGP 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          801
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                802 ATGAPTPPPAPPSP----SAPPPVVPKEEKEETAAAPPVEEGEEQKPPAAEELAVDTG 856
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   232 PSGSLAPPHHPMQPVSVNRQMNPANFPQLQQQQQQQQQQQQQQQQQQQQQQQQQQQQLQARPP 291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QQHQQQQPQGIRPQFTAPTQVPVPPGWNQLPSGALQPPPAQGSLGTMTANQGWKKAPLP- 350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRANASPQKPLDLKQLKQRAAAI 973
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9
                                -!- DOMAIN: Contains two Leu-Xaa-Xea-Leu-Leu (LXXLL) motifs. Only motif 1 is essential for the association with nuclear receptors, while adjacent Ser-884 displays selectivity for nuclear receptors.-!- PTM: Phosphorylated by PRKDC.
-!- PTM: Phosphorylation on Ser-884 leads to a strong decrease in
                                                                                                                                                                                   breast and lung cancers. CAUTION: Ref.1 (AAF16403) sequence differs from that shown due to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  172 GIIRMNNPATVMIPPGGNVSSSMMAPGPNPELQPRTPRPASQSDAMDPLLSGLHIQQQSH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     591 NSEEAITPQQSAELASMELNESSRWTEEEMETAKKGLLEHGRNWSAIARMVGSKTVSQCK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8 NLEDIYTSLCSSTMEDSEMDFDSGL--EDDDTKSDSILEDSTIFVAFKGNIDDKD----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      651 NFYFNYKKRONLDEIL------QQHKLKMEKER--NARRKKKKAPAASEEAAFPPV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              700 VEDEEMEASGVSGNEEEMVEEAEALHASGNEVPRGECSGPATVNNSSDTESIPSPHTEAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            857 KAEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTA---KSSGAPQ
prostate, testis and ovary; weakly expressed in lung, thymus and
                                                                                                                                          binding to ESR1 and ESR2.
MISCELLANEOUS: Frequently amplified or ovexpressed in colon,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity 18.9%; Pred. No. 6e-07;
Matches 408; Conservative 269; Mismatches 837; Indels 646;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             70. GO: 0005634; C:nucleus; IDA.
GO: GO: 0005667; C:transcription factor complex; TAS.
GO: 0003682; F:chromatin binding; ISS.
GO: 00030331; F:estrogen receptor binding; TAS.
GO: 00030331; F:restrogen receptor binding; TAS.
GO: 0003031; F:restroid hormone receptor binding; IDA.
GO: 000313; F:transcription co-activator activity; IDA.
GO: 0001031; F:transcriptional activator activity; IDA.
GO: 0001031; F:transcriptional activator activity; TAS.
GO: 0001031; F:hemarin development; ISS.
GO: 0001031; F:hemarin development; ISS.
GO: 0001031; F:hemari development; ISS.
GO: 000103099; P:hemari development; ISS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (sensu Mammalia); ISS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        760 KDTGONGPK----PP----ATLGADGPPP--GPPTPPRRTSRAPIEP---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3.2%; Score 428; DB 1; Length 2063; 18.9%; Pred. No. 6e-07;
                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; AF177388; AAF13595.1; -.
EMBL; AF208227; AAF16403.1; ALT_FRAME.
EMBL; AF245115; AAF74800.1; -.
EMBL; AF120468; AAF37003.1; -.
EMBL; AF171667; AAF71829.1; -.
EMBL; AL109824; CAB97231.1; -.
EMBL; AL109824; CAB97231.1; -.
                                                                                                                                                                                                                                 a frameshift in position 88.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                       This S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ---
                                                                                                                                                                                                           +
셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
```

¢.

_	1944 RVARPERPRADIGHAFLAKP	UD 1255 QINIFLEFERENLINGGFDQQCHNFILLEF		2064 LTHDKGLPKHLEELDK	1330	OV 2124 PGVKGHORVVTLAQHISEVITQDYTRHHPQQI	1355	Qy 2184 YLPPDHGAPARGS-PHSEGGKRSPEPNKTS	Db 1376 VSPTPLANPPVPGSFPNNSGLNPQNSTVS	Qy 2229EGMTEPGHSRSAVYPLLYRDGEQ:	Db 1434 KEQVNIELKAVPAQEVKMVVPEDQSKKDGQP	Qy 2281 AMVKSKKQEINKKLNTHNRNEPEYNISQP	DD 1478 NLVSPAMREAPTSLSQLLDNSGAPNVTIKPP	Qy 2331 AVQEHASTNMGLEAIIRKALMGKYI	DD 1538 SSSKEPSNSLNLPHSNELCSSLVHPELSEVS	Qy 2366 NAPNPLNASASLPAAMPITAADGRSDHT	Db 1598 TVFVTSNPITTSANTSAALPTHLQSALMSTV	Qy 2410SSRKAKSPAPGLASG		Qy 2450 WEDRPSSAGSTPFPYNPLIMRLQAGVM	Db 1715PNALSSSPAPNIQTGRPLVLSSRATPV		SULT 22 PB_MOUSE		DT 01-APR-1990 (Rel. 14, Created) DT 01-APR-1990 (Rel. 14, Last sequence upda				OC Mammalia, Eutheria, Rodentia, Sciurogr OX NCBL TaxID=10090;		STRAIN=Sw MEDLINE=9	RA Noble M., Lewis S.A., Cowan N.J.; RT "The microtubule binding domain of mic	MAP1	. : :	sphorylated MAP1B may accompany neurite e	two t
3513PMOOOLOARPSLAT365	PPIGVTKVHEPPREDAAPTKPAPPAPPPONLQPESDAPQQ-PGSSPRGKSRSPAPPADK 10	366VQT	1033 BAFAABAQKLPGDPPCWTSGLPFPVPPREVIKASPHAPDPSAFSY-APPGHPLPLGLHDT 1091	403OMKSLOGGPSRVPTPLOQPHLTNKSP-ASSPSSFQGSPASSPTVNQT 449	1092 ARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEHAKAPV 1145	450 QQQMGPRPPQ-NNPLPQGFQQPVSSPGRNPNVQQGNVPPNFWVMQQQDPNQ 499	GPVTMGLPLPMDPKKLAP-FSGVKQEQLSPRGQAGPPBSLG-VPTAQEASVLRGTALGSV		PGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPSRLDRGREDSL	PGNSGAPQLQANQNVQHAGGQGGAGFPQNQMQVSHGFPNNM	PKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETAAPKKIYDMMEGKVGKALS	QPSLMGIHGNMNNQQAGTSGVPQVNLSNMQGQPQQBP	SASIEGELMGKAIPPEKHSPHEUGEOHINGSIIQQIPKSIVEAGEDILKKEARULAN ::::::::::::::::::::::::::::::::::	-SQLMGMHQQ1VPSQGQMVQQQG1LNPQNPM1LSRAQLMPQGQMMVNPPS	BGTPPPPPPSRDLTEAYKTQALGPLKLKPAHEGLVATVKEAGRSI		HEIPREELRHTPELPLAPRPLKEGSITQGTPLKYDTGASTTGSKKHDVRSLIGSPGRT			788 LRPPGPSPHMAQQHGDPATTANNDVSLSQMMPDVSIQQTNMVPPHVQA 835	-APFAGHLPRGSPVTMREP-TPRLQEGSL	MQGNSASGNHFSGHGMSFNAPFSGAPNGNQMSCGQNPGFPVNKDVTLTSPLLVNLLQSDI		896 SAGHFGVNNKQNNTNANKPKKKKPPRKKKNSQQDLNTPDTRPAGLEEADQPPLPGEQGIS 955	LYRSHIPLAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYP-PYLIR	LDNSG-PKLPEFSNRPPGYPSQPVEQRPLQQMPPQLMQHVAPPPQPPQQPPQDPQCPQQQ	GYPDTAALENRQTIINDYITSQQMHHNTATAMAQRADMLRG-LSPRESSLALNYAA	1015 PPPPSQPQQQQQQQQQMMMMMQQDPKSVRLPVSQNVHPPRGPLNPDSQRMPMQQSG 1074	1727 GPRGIIDLSQVPHLPVLVPPTPGTPATAMDRLAYLPTAPQ-PFSS 1770	1075 SVPVMVSLQGPASVPPSPDKQRMPMVNTPLGSNSRKMVYQESPQNPSSSPLAEMA 1130	1771RHSSSPLSPGGPTHLTKPTTTSSSERERDRDRERDREREKSILTSTTTVEHAP 1825	1131 SLPEASGSEAPSVPGGPNNMPSHV	.1826 IWRPGTEQSSGSSGSSGGGGGSSSRPASHSHAHQHSP-ISPRTQDALQQRPSVLHNTGMK 1884	SHGHHFPN	PAATFPPATHCPLGGTLDGVYPTLMEPVLLPKEAP	1219 PYYPQTPNNRPPSTEPSEISLSPERLAASIAGLFPP 1254
-	à	QD	ò	qa	ò	qq	8	qq	ð i	gg .	λ _ο	g (È i	Q O	8	a	ò	q	δ	q	ò	đ	ò	qq	δλ	qq	λ	q	ò	ф	ò	qq	ò	qq	ò	Q

:: | : | 1375 OTEPSRMGSKSPGNTSQPPAFFSKLTESNS 2280 PG-TEI-FNMPAITGTGLM-----TYRSQ 2330 TLTSPGGGGKAKVS-----GRP---- 2409 GDRPPSVSSV---HSEGDCN-RRTPLTNRV 2449 LKAIGOAPSNLTMNPSNFAT----- 1299 ------SGKQSNSGATKRASPSNS-- 1329 OLSAPLPAPLYSFPGASCPVLDLRRPPSDL 2183 SVLGGG---EDGIEPVSPP----- 2228 -----PPLSA 2365 SSNVAPŠIPPVMSRPVSSSSISTPLPPNQI 1597 QELELRSLGYHGSSYSPEGVEPVSPVSSPS 2063 VKLGGEAAHLPHLRPLPESQPSSSPLLQTA 2123 SPSKGSEPRPLVPPVSGHATIARTPAKNLA 2003 PS is essentially unknown.
les in the cytoskeletal changes
les subly MABIB Binds to at least
r, and this bridging of subunits rotubule-associated protein if unrelated to that of MAP2 Vertebrata; Euteleostomi; thi; Muridae; Murinae; Mus. date)
update)
1B) (MAP1.2) (MAP1(X)) YDQWEES-----4 AA.

```
s SWISS-PROT entry is copyright. It is produced through a collaboration ween the Swiss Institute of Bioinformatics and the EMBL outstation - European Bioinformatics Institute.

Buropean Bioinformatics Institute of Buropean its on restrictions on its by non-profit institutions as long as its content is in no way
                                                                                                                                                                                                                                                                                              use by non-profit institutions as long as its content is in no way modified and this stetement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              314; Mismatches 996; Indels 869; Gaps 114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                68 LLSEFQPGNERSQELHLRPESHSYLPELG-----KSEMEFIESKRPRLELLPDPLL 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         119 RPSPLLATGQPAGSEDLTKDRSLTGKLEPVSPPSPPHTDPELELVPP-----RLSKE-- 170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----ELIQNMDRVDREITMVEQQI----- 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      263 YIFPGGRGDSALFAVNGFNMLINGGSERKSCFWKLIRHLDRVD---SILLTHIGDDNLPG 319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        --SKLKKKQQQLEEEAAKPPEPEKPVSPPPIESKHRSLVQIIYDENRKKAEAAHRILEGL 248
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -----RSQGSTSNSDWMKNLI 350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 249 GPQVELPLYNOPSDTRQYHENIKINQAMRKKLILYFKRRNHARKQWKQKFCQRYDQLMEA 308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | : : | | : : | | : | | 351 SPDIGVVFLAVPENLKDPEPLIKMKRSIEEACFTL------QYLNKLSMKPEPLFRS 401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          309 LEKKVE-----RIENNPRRRAKESKVREYYEKQFPEIRKQR-ELQERMQSRVGQ 356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Stabilizing microtubules.
SUBUNIT: 3 different light chains, LC1, LC2 and LC3, can associate with MAPIA and MAPIB proteins.
DOMAIN: Has a highly basic region with many copies of the sequence KKEE and KKEI/V, repeated but not at fixed intervals, which is responsible for the binding of MAPIB to microtubules.
FITH: LC1 is coexpressed with MAPIB. It is a polypeptide generated from MAPIB by proteolytic processing. It is free to associate with both MAPIA and MAPIB. It interacts with the amino-terminal region
 might be involved in nucleating microtubule polymerization and in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     174 LLSTTHPANKASLTL-----FCPEEGDWKNSNLDRHNLQDFINIKLNSASILP---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MAPIB 1.

MAPIB 2.

MAPIB 3.

MAPIB 3.

MAPIB 4.

MAPIB 5.

MAPIB 6.

MAPIB 6.

MAPIB 9.

MAPIB 9.

MAPIB 9.

MAPIB 10.

LYS-RICH (HIGHLY BASIC, CONTAINS MANY KKEE AND KKEI/Y REPEATS).

KKEE AND KKEI/Y REPEATS).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB 1; Length 2464;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Microtubule, Repeat, Phosphorylation.
CHAIN 7 2464 MAP1 LIGHT CHAIN LC1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 424.5; DB 1
Pred. No. 9.2e-07;
                                                                                                                                                                                                                                                                                                                                                                                                        PIR; S07549; QRMSP1.
MGD; MG1:1306778; Mtaplb.
OGC GO:0016358; P:dendrite morphogenesis; IMP.
GO; GO:0001578; P:microtubule bundling; IMP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | |::| :||||
320 INSMLQRKIAELEEE--------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro, IPR000102; MAP1B_neuraxin.
Pfam; PF00414; MAP1B_neuraxin; 10.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2464 AA; 270408 MW;
                                                                                                                                                                                                                                                                                                                                                                                       EMBL; X51396; CAA35761.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          19.3%;
                                                                                                                                                                                  of MAP1B.
SIMILARITY: TO MAP1A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches 521; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2010
2027
2044
589
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ठ
```

충 음 ð	357 RGSGL-SMSAARSEHEVS	RGSGL-SMSAARSEHEVS 394
} 8	TGQVPTPPVKQVKLKQR	_
රු අ	449 KNFGLIASFLERKTVAECVLYYYLTKKNE	KNFGLIASFLERKTVAECVLYYYLITKKNENYKSLVRRSYRRGKSOQQQQ0QQQQQQO 506
δ d	507 QQQQPMPRSSQBEKDEKE-KEKEAEKEEE	QQQQPMPRSSQBEKDEKE-KEKEABKEEEKPEVENDKEDLLKEKTDDTSGEDNDEKEAVA 565
ò 8	566 SKGRKTANSOGRR	SKGRKTANSQGRRKGRITRSMANBANSEBAITPQQSAELASMELN 610 ::
è a	611 ESSRWTEEEMETAKKGLLEHGRNWSAIAR 	ESSRWTEBEMETAKKGLLEHGRNWSAIARWVGSKTVSQCKNFYFNYKKRQNLDEILQOHK 670
Qy Op	671 LKWEKERNARRKKKKAPAAASEEAAFPPVVEDEEMEASG 	-apaaseeaafppvvedeemeasgysgneeemveeaealha 726
& g	727 SGNEVPRGECSGPATVNNSSDTESIPSPHTEAAKDT	SGNEVPRGECSGPATVNNSSDTESIPSPHTEAAKDTGQNGPKPPATLGADGPPPGPPPP 786
& A	787 RRTSRAPIEPTPASEATGAPTPPPAPPSPSAPPPVVPKEEKEËETAAAPPVB 	SAPPPVVPKEEKEETAAAPPVBEGE 841
& g	842 BQKPPAABBLAVDTGKABBPVKSECTBBA : : : 900 GECEQTPBELBPVBKQGVDDI	EQXPPAAEELAVDTGKAEEPVKSECTEEAEEGPAKCKDAEAAEATAEGALKAEKKEG 898
& q	899GSGRATTAKSSGAPQDSDSSATCSADE 	GSGRATTAKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRANA 955
& A	SPQKPLDLKQLKQRAAAIPP : : : DMDDVLEKGEAEOS	IOVTKVHEPPREDAAPTKPAPPAPPPQNLQPESD 1010 -EREGEEDKAEDAREGYEPDKTBAEDYWAVADKAABA 1052
ර් යි	GVTEEQYGYLGTSAKQPGIQ	ADKEAFAAEAQKLPGDPPC 104 :: :: 110 IHDETLPGGSESEATASDEENREDQPEE 110
& <u>8</u>	1049 WTSGLPPPVPPREVIKASPHAPDPSAFSYAPPGHPLPLGLHDTARPVLP 	APPGHPLPLGLHDTARPVLPRPPTISN 1104
දු පු	ERQ YSQEYSKPAVA	-IGAISQGMSVQLHVPYSEHAKAPVGPVTWGLPLP 1155
දු පු	1156 MDPKKLAPFSGVKQEOLSP 	-PFSGVKQEQLSPRGQAGPPESLGVPTAQEASV 1194
oy Op	1195 LRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTP :	LRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPSR 1254
්ර යි	1255 LDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGP1309 EEKTLEVVSPSQSVTGSAGHTPXYQSPTDEKSSHLPTE	MSVTQCSKEDGRSSGRPHETAAPKRT 1310
ò	1311 YDMMEGRVGRAISSASIEGLMGRAIP	YDWMEGRVGRAISSASIEGLMGRAIPPERHSPHHLKEQHHIRGSITQGIPRSYVE 1365

		Db 2252 SPVKKGDGK-
q	1356VSFEFSEAKDENERASLSPMDEPVPDSESPVE-KVLSPLRSPPLLGSESPY 1405	2302
ò	1366 AQEDYLRREAKLIKREGTPPPPPPSRDLTEAYKTQALGPLKLKPAHEG 1413	 Db 2302 PEV
q	1406EDFLSADSKVLGRRSESPPEGKNGKQGFPDRESPVSDLTSTGLYQDKQEEKSTG 1459	Qy 2362 PLSANAFNPLNAS
ò		:
q		:
δ		SULT 23 T1_YEAST
q		ID SNT1 YEAST STANDAR AC P25357; Q02397; Q8NIL
ζ	1534 GKPRQSPLTYEDHGAPFAGHLPRGSPVTMREPTPRLQBGSLSSSKASQDRKLTSTPREI- 1592	DT 01-MAY-1992 (REI. 22, DT 10-OCT-2003 (REI. 42,
qa	1562 TTDDVSPSLHAEVGSPHSTEVDDSLSVSVVQTPTTFQETENSPSKEECPRPNSISPDFS 1621	
ζŏ	1593 AKSPHSTVPEHHPHPISPYEHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLDAAAAY 1649	
පු	1622 PKTAKS-RTPVQDHRSE05SMSIEFGQESPEHSFAMDFS 1659	
ò	1650 YLPRHLAPNPTYPHLYPPYLLRGYPDTAALENRQTIINDYITSQQMHHNTATAMAQRADM 1709	
qa		
ò	1710 LRGLSPRESSLALNYAAGPRGIIDLSQVPHLPVLVPPTFGTPATAMDRLAYLPTAPQ 1766	
q	1686SPCDIQDSSLSHK	
ð	1767PFSSRHSSSPLSPGGPTHLTKPTTTSSSERERDRDREREK 1812	[2] SEQUENCE FROM N.A.
g	1722 VEASPETSSAHTPSQIASPLOEDTLSDVVPPREMŠLYASLASEKVOSLEGEKLSPK 1777	
ŏ	PGTEQSSC	
qq	:: : :	10.1 Yeast
ò	HNTGMKGI ITAVEPSKPTVLRST	
q		RC STRAIN=S288C; RX MEDLINE=92244356; Pub
ò	1933 MEPVILLPKEALAKPPAR 1966	
đ	1872 ENAAGSPDEEDYDYESQEKTIRTHDVVRYYYEKTERTIKSPCDSGYSYETIEKTTKTPED 1931	
ò	1967 SGLEPASSPSKGSEPRPLVPPVSGHATIARTPAKNLAPH 2005	
g	1932 GGYTCEITEKTIRIPEEGGYSYELSEKTTRIPEVSGYTYEKTERSKRLLDDISNGYDDTE 1991	
ò	2006HASPDPPAPPASASDPHREKTQSKPFSIQELELRSLGYHGSSYSPEG 2052	
: A	1992 DGGHTLGDCSYSYETTEKITSFPBSESYSYETSTKTTRSPDTSAYCYET 2040	
ò	2053 VEPVSPVSSPSLTHDKGLPKHLEELDKSHLEGELRPKOP 2091	RA Hollenberg C.P., Huse
; <u>8</u>	: :	
è		
7 E		RA Newlon C.S., Olson M
ìè	RHHPOOLSAPLPAP	
5 A		
è	2193 PARGSPHSEGGKRSPEPNKTSVLGGGEDGIEPVSPPEGWTEPGHSRSAV 2241	RA Spiegelberg R., Stat
2 40		RA van Vliet-Reedijk RA von Wettstein D.,
ò	VPLLYRDGEQ	RA Zimmermann F.K., Sgo RT "The complete DNA se
•		

2301	2361	2313	2421	2350																														٠.						
	KYDQWEESP -		SSRKAKSPAPGL	2314 KETKNAANASASKSAKTATTGPGTTKTAKSSTVPPGL 23	YEAST SNT1_YEAST STANDARD; PRT; 1226 AA.	P25357; Q02397; Q8NIL8; 01-MAY-1992 (Rel. 22, Created)	10-OCT-2003 (Rel. 42, Last sequence update) 10-OCT-2003 (Rel. 42, Last annotation update)	Potential DNA-binding protein SNT1. SNT1 OR YCR033W OR YCR33W OR YCR592.	ron	charomycetales; I_TaxID=4932;	[1] SEQUENCE FROM N.A.	, Brown A.J.P.;	The complete sequence of a 7.5 kb region of chromosome III from sarcharomores carevisiae that lies between CRY1 and MAT.";	Yeast 7:761-772(1991).	[2] SECITEMOR PROM N.A.		"The complete sequence of the unit YCR59, situated between CRY1 and "The complete sequence of the unit frames, which cover 91% of the		Yeast 7:413-424(1991).	131 SEQUENCE FROM N.A.		Oliver S.G., van der Aart Q.J.M., Agostoni-Carbone M.L., Algle M.,	Alberghina L., Alexandraki D., Antoine G., Anwar K., Bariesta C Renit P. Berben G., Bergantino E., Biteau N., Bolle PA.,	m	the Carignani G., Chanet R., Contreras K., Crouzet M., Daighan, Committee C., Dubois E.,	Dujon B., Duesterhoeft A., Erdmann D., Esteban M., Fabre F.,	Fairnead C.A., Faye G., Ferumann 11., Firm I., Fukuhara H., Francinques-Gaillard MC., Francinques-Gaillard C., Francindues-Gaillard C., Francind	Fuller L.J., Gent M.E., Gigot D., Gilliquet V., Glansdorff N.,	Hatat D., Hegemann J.H., Herbert C.J., Hilger F., Hohmann S.,	. :	Kleinhans U., Kreisl P., Lafranchi G., Lewis C., van der Linden C.G.,	Incchini G., Lutzenkirchen K., Maat C., Mainnaupt G., Mariem, Martegani R., Mathieu A., Maurer C.T.C., McConnell D., McKee R.A.,	Messenguy F., Mewes HW., Molemans F., Montague M.A., Navas L.,	Newlon C.S., Olson M.V., Faillet C., Planta R.J., Plevani P., Perea J., Philippsen P., Pierard A., Planta R.J., Possing C. W.	Poetsch B., Pohl F.M., Purnelle B., Ramezanı kad M., Kamussen B.,.,	Sanz E., Schaaff-Gerstenschlaeger I., Scherens B., Schweitzer B.,	Shu Y., Skala J., Slonimski P.P., Sor F., Sousceile C., Chierry A., chierelhero R., Thierry A.,	Thireos G., Triano L.N., Urrestarazu L.A., Valle G., Vetter I.,	van Viet-Reedlik J.C., Volckaert G., Viekon F., Murst H., Xu G., von Wettstein D., Wicksteed B.L., Wilson C., Wurst H., Xu G.,	<pre>Zimmermann F.K., Sgouros J.G.; "The complete DNA sequence of yeast chromosome III.";</pre>
q	ò	QQ	δλ	g	io H	AC				88	R R	XX e	Z Z Z	RL	RN	2 Z I	K E	R F	R.	R P	2 X	RA	RA Pa	§ §	RA RA	\$	RA RA	Æ	8 8	RA	8 8	RA 0	2 2	a a	RA S	\$ \$	Z a	8 8 8	& &	RA

```
| : : | | | : : | | | : | | | | PLDHIDTLADAASSVTNNONF---SNERNAIDIGRKSTTISNLLNNSDRSMKSSFQSASR 1113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           969 RPNINAYSNIPPQQRPALGYFVGQPTHGHNTSISSIDGSIRPFGPDFH----RDTFSKIS 1024
                                                                                                                                    DLTKDRSLTGKLEP----VSPPSPP----158
                                                                                                                                                                                                                                                                                                                                                       DETKAISYHDNNFPPRKSVSKPNAPLEPDNIKVGEEDALGKKEVHKSGREIAKEHPTPVK 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          QRYDQLMEALEKKVE---RIENNPRRRAKESKVREYYEKQFPEIRKQRELQERMQSRVGQ 356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AECVLYYYLTKKNENYKSLVRRSYRRRGKSQQQQQQQQQQQQQQQQQQQPMPRSSQEEKDEK 523
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EECVLHYYRTKKTVNYKQLLIDKNKKRKMSAAAKRRKKERSNDEEVEVDESKEESTNTI 765
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----ORNAARNGWKL-----LVDETDLKRDGTS---SESVQQSQILIQPE 968
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 883 AEATAEGALKAEKKEGGSGRATTAKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRP 942
----HSRDYASHLSPGSIIOPOR 63
                                                                                                                                                                                                                                                                                                                                                                                                                                         ------YHENIKINQ----AMRKKLILYFKRRHARKQWKQKFC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | : | : | : | : | 1: | 4.8 KIISEQKYLLKKAIRNFSEYPFYAQNKLIHQQATGLILTKIISKIKKEEHLK---KINLK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                535 HDYFDLOKKYEKECEILTKLSENLRKEBIENKRKEH-----ELMEOKRREEGIETEKEK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     357 RGSGLSMSAARSEHEVSEIIDGLSEQENL-----EKQMRQLAVIPPMLYDADQQR-IK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SLRHPSSSSSSRRNRADFVDD-AEMENVLLQIDPNYKHYQAAATIPPLILDPIRKHSYK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    886 EHKISYWSVRESQLFPELLKEFGSOWSLISEKLGTKSTTWVRNYY------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ----HASGNEVPR-----GECSGPATVNNSSDTESI-----PSPHTEAAKDTGQNGP
                                               PESMNTST-ITHTNTDIGNSRYSRKTMSRYNPQSTSSTNVTHFPSALSNAPPFYVANGSS
                                                                                                                                                                                                                                                  241 TITSNINSIHQREPFWKANSTILLKSIHSQSSPSLHIKKFHDANKLDKPEASVKVEIPSK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             --EKPVSPPPIESKHRSLVQIIYDENRKKAEAAHRILEGL---GPQVELPLYNQPSDTRQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FINANGLMAD----PMKVYKDRQVMNMWSEQEKETFREKFMQHPKNFGLIASFL-ERKTV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EKEKEAEKEEEK------DEVENDK------BDLLKEKTDDTSGE-----DNDEK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EAVASKGRKTANSOGRRKGRITRSMANEANSEAITPOOSAEL------ASME
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HKLIOMEKERNARRKKKKKAPAAASEEAAFPPVVEDEEMEASGVSGNEEEMVEEAEAL---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           768 KPPATLGADGPPPGPPTPPRRTSRAPIEPTPASEATGAPTPPPAPPSPSAPPPVV----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----MAEPTVTDLRNR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PKEEKEEETAAAPPVEEGEEOKPPAAEELAVDTGKAEEPVKSECTEEAEEGPAKGKDAEA
                                                                                                                                                                                                                                                                                                                                                                                                            ----ELELVPPRLSKEELIQNMDRVDREITMVEQQISKLKKKQQQLEEEAAKPPEP---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LNESSRWTEEEMETAKKGLLEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDEILQQ
                                                                                                                                                                                                   ---SEMEFIESKRP----RLELLPDPLLRPSPLLATGQ----
                                                                                                 RRPSLLSEFQPG-----NERSQELHLRPESHSYLPELGK-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1025 APLTTL --- PPP------RLPSIQFPRSE
PHSLSYPVQIARTHTDVGLLEYQH----
                                                                                                                                                                                                                                                                                                                                                         301 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      648
21
                                                                                                 64
                                                                                                                                                                                                   98
                                                                                                                                                                                                                                                                                                                                                                                                            159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       589
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           90/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             524
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   609
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       931
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               823
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                    셤
                                                                                               ò
                                                                                                                                      셤
                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                      ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            44;
                                                                                                                                                                                              MEDLINE=21567937; PubMed=11711434;
Pijnappel W.W.M.P., Schaft D., Roguev A., Shevchenko A., Tekotte H.,
Wilm M., Rigaut G., Seraphin B., Aasland R., Stewart A.F.;
"The S. cerevisiae SET3 complex includes two histone deacetylases,
Hos2 and Hstl, and is a meiotic-specific repressor of the sporulation
gene program.";
                                                                                                                                                                                                                                                                                                                          gene program.";
Genes Dev. 15:2991-3004(2001).
-!- FUNCTION: Part of the Set2C complex, which is required to repress
early/middle sporulation genes during meiosis.
-!- SUBUNIT: Identified in a Set3C complex with SET3, HST1, HOS2,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Germonline; 138939; -.
SGD; S0000629; SNT1.
GO; GO:000118; Cinistone deacetylase complex; IDA.
GO; GO:000118; Cinistone deacetylase complex; IDA.
GO; GO:0017136; F:DAD binding; NAS.
GO; GO:0017136; F:NAD-dependent histone deacetylase activity; IDA.
GO; GO:0015129; F:NAD-independent histone deacetylase activity; IDA.
GO; GO:0016575; P:nhistone deacetylation; IDA.
GO; GO:0016575; P:negative regulation of melosis; IDA.
InterPro; IPR001015; Myb_DNA_binding.
                                                                                                                                                 IDENTIFICATION IN A COMPLEX WITH SET3; HST1; HOS2; SIF2; CPR1 AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels 343;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DB 1; Length 1226;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  70AE6622227088EB CRC64;
                                                                       Valles G., Volckaerts G.;
Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SER-RICH.
COILED COIL (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ?
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 3.1%; Score 415.5; DB 1; Best Local Similarity 19.7%; Pred. No. 9.3e-07; Matches 236; Conservative 176; Mismatches 440;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (IN REF. 2)
(IN REF. 2)
(IN REF. 2)
(IN REF. 1)
(IN REF. 2)
                                                                                                                                                                                                                                                                                                                                                                                                                                                        SIF2, CPR1 and HOS4.
SUBCELLULAR LOCATION: Nuclear (Probable).
SIMILARITY: Contains 1 Myb-like domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NI)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POLY-SER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           H X H 4 O X > > O X H &
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -> A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AVQ ->
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        QQ < Q B H L B Z < Q Q P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Σ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL, S78624, AAB21259.1; -. EMBL, X59075, CAA41799.1; -. EMBL, X59720; CAC42983.1; -. PIR; S15053, S15053.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     138396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           934
599
305
375
4404
442
513
Nature 357:38-46(1992)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      888
3305
4405
4405
4405
7466
831
831
                                                  REVISIONS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
DNA_BIND
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CON
```

534

210

463

```
TISSUE-12388257; PubMed=12477932; Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Strausberg R.L., Feingold E.A., Grouse L.H., Schaefer C.R., Schuler G.D., Altechul S.F., Zeeberg B., Batcow K.H., Schaefer C.R., Bhat N.K., Altechul S.F., Zeeberg B., Batcow K.H., Schaefer C.R., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haich F., Antechul S.F., Zeavant T.L., Schaefer T.E., Stapleton M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunztane P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Milalon D.K., Muzry D.M., Sodergren E.J., Lu X., Gibbs R.A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E., Schmutz J., Dickson M.C., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E., Schmutz J., Marra M.A., Schnerch A., Schnerch A., Schmutz J., Marra M.A., Schnerch A., Schnutz J., Schnutz J., Marra M.A., Schnutz J., Sch
                      1114 HEAQLE------KPNITTPRS 1141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    "Aczonin, a 550-kd putative scaffolding protein of presynaptic active cones, shares homology regions with rim and bassoon and binds profilin.";
                                                                                                                        Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.; Prediction of the coding sequences of unidentified human genes. IX. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro.";
                                                                                   SLLTPTGDPRANASPOKPLDLKQLKQRAAAIPPIQVTKVHEPPREDAAPTKPAPP 997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Wang X., Kibschull M., Laue M.M., Lichte B., Petrasch-Parwez E., Kilimann M.W.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SECUENCE OF 4405-5147 FROM N.A. Kalicki J., Elliott G.; Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SECUENCE OF 552-4404 FROM N.A.
Kraemer J., Wollam C., Wohldmann P., McGrane B.;
Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002)
                                                                                                                                                                                                                                                                          PCLO HUMAN STANDARD; PRT; 5147 AA. 094570; 043373; 060305; 09BVC8; Q9UIV2; Q9Y6U9; 28-FEB-2003 (Rel. 41, Created) 28-FEB-2003 (Rel. 41, Last sequence update) 10-OCT-2003 (Rel. 42, Last annotation update) PICCO.D protein (Aczonin) (Fragments).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE OF 3619-5147 FROM N.A. (ISOFORM 2)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IISSUE=Brain;
MEDLINE=99439764; PubMed=10508862;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE=98290545; PubMed=9628581;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE OF 4405-4439 FROM N.A. TISSUE-Placenta;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Cell Biol. 147:151-162(1999)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE OF 1-759 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TISSUE=Brain;
                                                                                           943
                                                                                                                                                                                                                                                                                     REPAREMENT REPAREMENT OF THE PROPERTY OF THE P
                                     용
                                                                                                                                            셤
                                                                                           ð
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      C2 DOMAIN 1.
C2 DOMAIN 2.
C2 DOMAIN 2.
S -> SGNGLGIRIVGGKEIPGHSGEIGAYIAKILPGGSAE
OTGKLMEG (in isoform 2).
//FTIda-VSP 003923.
K -> KPTDGTKVVSHPITGEIQ (in isoform 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              10 X 10 AA TANDEM APPROXIMATE REPEATS OF P-A-K-P-Q-P-C-X.
C4-TYPE (POTENTIAL).
C4-TYPE (POTENTIAL).
                                                                                                                                                                                                                                                                                     Note=No experimental confirmation available;

VSP_003924, VSP_003924, VSP_003925, VSP_003927;

I DOMAIN: C2 domain 1 is involved in binding calcium and phospholipids. Calcium binds with low affinity but with high specificity and induces a large conformational change.

SIMILARITY: Contains 2 C2 domains.

SIMILARITY: Contains 1 PDZ/DHR domain.
-i. FUNCTION: May act as a scaffolding protein involved in the organization of synaptic active zones and in synaptic vesicle trafficking (By similarity).
-i. SUBUNI: Interacts with Rabacl/Pral, RIMS2 and profilin (By similarity).
-i. SUBCELLULAR LOCATION: Concentrated at the presynaptic side of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GO; GO:0005856; C:cytoskeleton; NAS.
GO; GO:0005856; C:cytoskeleton; NAS.
GO; GO:0005502; C:synaptic junction; ISS.
GO; GO:0005549; F:calcium ion binding; ISS.
GO; GO:0005522; F:profilin binding; ISS.
GO; GO:0007010; P:cytoskeleton organization and biogenesis; ISS.
GO; GO:0016080; P:synaptic vesicle targeting; ISS.
InterPro; IPR000008; C2.
InterPro; IPR001565; Synaptotagmin.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRINTS; PRO0399; SYNAPTOTAGMN.
PRINTS; PR00399; SYNAPTOTAGMN.
SMART; SM00239; C2; 2.
PROSITE; PS500049; C2_DOMAIN 1; 1.
PROSITE; PS500049; C2_DOMAIN 2; 2.
Calcium/phospholipid-binding; Zinc; Metal-binding; Zinc-finger;
Repeat; Alternative splicing.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /FTId=VSP_003924.
G -> GQVMVVQNAS (in isoform 2).
/FTId=VSP_003925.
                                                                                                                                                                                                    Bvent=Alternative splicing, Named isoforms=2;
Comment=Additional isoforms seem to exist;
                                                                                                                                                                                                                                                                              IsoId=Q9Y6V0-1; Sequence=Displayed;
                                                                                                                                                 synaptic junctions (By similarity).
ALTERNATIVE PRODUCTS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POLY-PRO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; ACC04903; AAD20936.1; --
EMBL; ACC04886; AAD21789.1; --
EMBL; AB011131; BAA25485.1; --
EMBL; ACC01104; AAH01304.1; --
EMBL; ACC04082; AAB97937.1; --
PIR; TO0634; TO0634.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; Y19188; CAB60727.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Genew; HGNC:13406; PCLO.
MIM; 604918; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        523
992
1011
2325
4442
4633
5121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   4534
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 4576
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4534
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 4576
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        499
969
1010
2300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4391
4544
5031
4404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           400
                                                                                                                                                                                                                                                            Name=1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ZN_FING
NON CONS
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
DOMAIN
VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ZN FING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
```

& & & & & & & & & & & & & & & & & & &	8 6 8 6	8 6 8 6	6 8 8 8	8 6 8 6	8 8 8 8 8 8	8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8
VARSPLIC 4757 4761 TAHKS -> SKRKK (in isoform 2). /FTIG=VSP 003926. VARSPLIC 4762 5147 Missing (In isoform 2). /FTIG=VSP 003927. SEQUENCE 5147 AA; 563537 MW; CD5D84990498CD3C CRC64; Query Match 3.1%; Score 410.5; DB 1; Length 5147; Best Local Similarity 19.0%; Pred. No. 5.1e-06; Matches 477; Conservative 276; Mismatches 895; Indels 863; Gaps 105;		AAFPEVVEDEEMEASGVSGNE	764 QNGPKPPATLGADGPPGPPTPPRRTSRA	### ### ##############################	970 AAAIP-PIQUTKVHEPPREDAAPTKPAPPAPPPPQNLQPESDAPQQPGSS-PRG 1021 394 PAKAPGPTKTPAQTKPPSQQPGSTKPPPQQPGPAKPSPQQPGSTKPPSQQPGSAKPSA 451 1022 KSRSPAPPADKEAFAARAQKLPGDPPCWTSGLPFPVPPREVIKASPHAPDPSAFSYAPPG 1081 1022 KSRSPAPPADKEAFAARAQKLPGDPPCWTSGLPFPVPPREVIKASPHAPDPSAFSYAPPG 1081 1082 HPLP	1122 IGAISQGMSVQLHVPYSEHAKAPVGPVTMGLPLPM
FT FT FT SQ Que Bes	8 8 8	8686	8 6 8 6	8 6 8 6	8 6 8 6 8 6	8 6 8 6 8 6 8 6 8 6

>-	1353 -	GSITOGIPKSYVEAQEDYLRREAKLLKREGTP 1384
۵	815 F	
>-	1385 -	PPPPPSRDLTEAYKTQALGPLKLKPAHEGLVATVKEAG 1422
۵	874 M	MKQAPAPSQPPTSQGPPKSTGQAPPAPAKSIPVKKETKAPAAEKLEPKAE-QAPTVKRTE 932
>-	1423 R	RSIHBIPKYDTGASTTGSK 1469
۵	933 T	EKKPPPIKDSKSLTAEP
>-	70	SPGR
۵	993 K	KNQVCNLCGFNPTPHLTENCQTQRAISGQLGDİRKMPPAPSGPKASPMPVPTE 1045
٠.		PFAGHLPRGSPVTMREPTPRLQE 157
ٔ م	6	PQVKLVKKQEQEVKTEAEKVILEKVKETLSMEKIPPMVTTDQKQEE 110
م ح	1572 G	GSLSSSKASGBRYUP 1601 GRISSSKAS
>-	1602 E	EHHPHPISPYEHLLRGVSGVDLYRSHIPLAPDPTSIPRGIPLDAAAAYYLPRHLAPNPTY 1661
۵	1161 B	
>-	1662 P	PHLYPPYLIRGYPDTAALENROTIINDYITSQOMHHNTATAMAQRADMLRGLSPRESS 1719
۵	1190 E	EGKOPOTKMEGLPSGTPQSLPKEDDKTTKTIKEQPQPPCTAKPDQEKEDDKSDT 1243
>-	1720 L	LALNYAAGPRGIIDLSQVPHLPVLVP
۵	1244 S	
>-	1752 -	
۵	1304 L	LAKLESTVLSILEAQASTLADEKSEKKTQPHEVSPEQPKDQEKTQSLSETLEITISEEEI 1363
>	-	VEHAPIWRPG 183
Ω	1364 K	KESQEERKDTFKKDSQQDIPSSKDHKEKSEFVDDITTRREPYDSVEESSESENSPVPQRK 1423
> 1		R 186
n	ď	EFIRNOLKEISSS 148
> 0	1867 T 	TODALQORPSVLHNTGMKG11TAVBESKPTVLRSTGTSSPVRPAATFPPATHCPLGGTLD
.	1927 G	PVLL
0	1514 -	: ::
>-	1987 P	ASDPHREKTQSKPFSIQELE
0	1553 -	-VTGGGGLRRFKTIELNSTIADKYSAESSQKKTSLYFDEEPELEMESL 1599
>-	2042 -	
0	1600 T	DSPEDRSRGEGSSSLHASSFTPGTSPTSVSSLDEDSDSSPSHKKGESK 1648
Sec. 0	α σ	PKOPGPVKLGGGAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQRVVTLAQHI 2139 :: :
:	2140	
	1697 S	SSSKKSKKDKDELRAQRRRERPKTP-PSNLSPIEDAS-PTEELRQAAEMEELHRSSC 1751

```
This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                 2096 TILDSITTVYTEPVDMITKFEDSEEISSSTYFPGSIIDYPEEISVSLDRTAPPDGRASAD 2155
                                                                                                                                                                                                                                                                    ---SPAPGLASGDRPPSVSSVHSEGDCNRRTPL 2445
                                                                                                                        1869 LTRQEEENGFMQQKGREQKI--RLSEQIYEDFMQKITDLQKEFYELESLHSVVPQEDIVS 1926
                                                                                                                                                                                  SSFIIPESHEIVDLGTMVTSTEBERKLLDADAAYEELMKRQQM------QLTPGSSPTQ 1979
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SUBMINIT: 3 different light chains, LCI, LC2 and LC3, can associate with MAPIA and MAPIB proteins.

DOMAIN: Has a highly basic region with many copies of the sequence KKEs and KKEI/V, repeated but not at fixed intervals, which is responsible for the binding of MAPIB to microtubules.

PTM: LCI is coexpressed with MAPIB. It is a polypeptide generated from MAPIB by proteolytic processing. It is free to associate with both MAPIB and MAPIB. It interacts with the amino-terminal region of MAPIB (By similarity).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-NOV-1995 (Rel. 32, Created)
01-NOV-1995 (Rel. 32, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Microtubule-associated protein 1B (MAP 1B) [Contains: MAP1 light chain
                                                                                                                                                                                                                                  SEYSPSIES-DPEGFEISPEKIIEVQKVYKLPTAVŠLYSPTDE--QSIMQKEGSQKALKS
                                                                                                                                                                                                              NPL-------NASASLPAAMPITAADGRSDHTLTSPGGGGKAKVSGRPSSRK
                                                                                                                                                                                                                                                                                              2036 IEDEYVTDYTREIQEIIAHESLILTYSEPSESATSVPPSDTPSLTSSVSSVCTTDSSSPI
                                                                                                                                                                                                                                                                                                                         TN------RVWEDRPSSAGSTPFPYNPLIM--RLQAGVMASPPPPG----
--- PNKTSVLGGGEDGIEPVSPPEGMTEPGHS
                                                     -GNTSOPPAF
                                                                            AEEMYEEMMHKTHKYKAFPAANERDEVFEKEPLYGGMLIEDYIYESLVEDTYNGSVDGSL
                                                                                                       FSKLTESNSAM-VKSKKQEINKKLNTHNRNEPEYNISQPGTEIFNMPA------
                                                                                                                                                          -----ITGTGLMTYRSQAVQEHASTNMGLEAIIRKALMGKYDQWEESPPLSANAF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Homo sapiens (Human).
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                              ----LPAGSGPLAGPHHAWDEEPKPLLCSQYETLSDSE 2517
                                                                                                                                                                                                                                                                                                                                                                                               2468 AA
                                                       RSAVYPLLYRDGEQ-----TEPSRMGSKSP-
     DHGAPARGSPHSEGGKRSPE-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STANDARD;
                                                                                                                                                                                                                                                                           2414 AK-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HUMAN
                                                                                                                                                                                                                                                                                                                                                                                    2484
                                                                                                                                                                                             1927
                                                                                                                                                                                                                        2369
                                                                                                                                                                                                                                                                                                                                 2446
                                                                                                               2272
                                                                                                                                                                    2319
                                                           2238
                                                                                    1809
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           P46821
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MAP1B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MAPB
                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT
MAPB_HU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                        ઠે
                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                            8 8
                                                                                                                                     셤
                                                                                                                                                                    ð
                                                                                                                                                                                         g
                                                                                                                                                                                                                       8
                          g
                                                          8
```

```
432
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           491
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          648
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HARKQWKQKFC----QRYDQLMEALEKKVERIENNPRRRAK-------ESKVREY 332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       119 RPSPLLATGQPAGSEDLTKORSLTGKLEPVSPPSPPHTDPELELVPP-----RLSKE-- 170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            263 YIFPGGRGDSALFAVNGFNMLINGGSERKSCFWKLIRHLDRVDSILLTHIGDDNLPGINS 322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   323 MLORKIAELEEEGSSGSTINSDWMKNLISPDLGVVFLNVPENLKNPEPNIKMKRSIEEAC 382
between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEIIDGLSEQENLEKOMROLAVI PPMLYDADQORIKFINMNGLMADPMKVYKDROVMNMM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQEKETFREKFMQHPKNFGLIASFLERKTVAECVLYYYLIKKNENYKSLVRRS-YRRRG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           KSQQQQQQQQQQQQQPMPRSSQEEKDEKE----KEAEKEEEKPBVEND----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        tgtnkdkarfilpnooevdlpisyltsvsslivwhpanpaekiirvlfpgnstoynileg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       YEX------OFP-EIRKORELOERMOSRVGORGSGLSMSAARSEHEV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  YDEN--RKKABAAHRILEGLGPQVELPLYNQPSDTRQYHENIKINQAMRKKLILYFKRRN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FTLQYLNKLSMKPEPLFRSVGNTIDPVILFOKMGVGKL-EMYVLNPVKSSKEMQYFMQQM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -----KDKPV----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----EMBGLSEFTEYLS----ESVEVPSPFDI-----LEPPTSGGFLKLSKPCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LLSEFQPGNERSQELHLRPESHSYLPELG-----KSEMEFIESKRPRLELLPDPLL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----ELIQNMDRVDREI-----T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MVEQQISKLKKKQQQ----LEBERAAK---PPE-----PEKPVSPPPIESKHRSLVQII
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CONTAINS MANY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3.1%; Score 410; DB 1; Length 2468;
llarity 17.8%; Pred. No. 2.6e-06;
Conservative 360; Mismatches 1021; Indels 882;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2468 AA; 270618 MW; 540839CBDF09D461 CRC64;
                                                                                                                                                                                                                                                                                                                                              MAP1 LIGHT CHAIN LC1.
MAP1B 1.
MAP1B 2.
MAP1B 3.
MAP1B 4.
MAP1B 5.
MAP1B 6.
MAP1B 6.
MAP1B 1.
MAP1B 9.
MAP1B 9.
MAP1B 9.
MAP1B 9.
MAP1B 9.
MAP1B 9.
                                                                                                                                                                                                                                   GO; GO: 0005875; C:microtubule associated complex; TAS.
InterPro; IPR00102; MAPIB neuraxin.
Prom; PF00414; MAPIB neuraxin; 10.
PROSITE; PS00230; MAPIB NEURAXIN; 6.
Microtubule; Repeat; Phosphorylation.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  562 KEETPEVTKVNHVEK------PPKVESKEKVMVK-
                                                                                                                                                                      EMBL; L06237; AAA18904.1; -. Genew; HGNC:6836; MAP1B.
                                                                                                                                                                                                                                                                                                                                                                                   1894
1911
1928
1945
1962
1979
2013
2047
2064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity
                                                                                                                                                                                                                                                                                                                                                                                      1878
1895
1912
                                                                                                                                                                                                                                                                                                                                                                                                                                                          1929
1946
1963
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1997
2014
2031
2048
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      589
                                                                                                                                                                                                                            157129; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 649
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    502
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                          REPEAT
REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                              REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REPEAT
                                                                                                                                                                                                                                                                                                                                                                                          REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8
                   SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOLUTION SOL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8
```

600	660 776	718 823	778	863 838	894	893 946	947	1007	1041	1101	1097	1143	1204	1181 1258	1241	1300	1301 1350	1361	1363	1414		1462	1495	1530
	11 SABLASMBLNESSSRWTEBEMETAKKGLLEHGRAMSAIARWYGSKTVSQCKNFYFNYKKRQ 	11 NLDEILQQHKLKMEKERNARRKKKKAPAASEEAAFPPVVEDEEMEASGVSGNEEEMV	EEA	824EAERSLMSSPEDLTKDFEELKAEEVDVTKDIKPQLELIED 779 PPGPPTPPRRTSRAPIEPTPASEATGAPTPPPAPPSPSPSAPPPVVPKEEKEEEFAAAPPVE		839EGEEGKPPAAEELAVDTGKAEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKA :	14 EKKEGGSGRATTAKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTP	TGDPRANASPQKPLDLKQLKQRAAAIPPIQVTKVHEPPREDAAPTKPAPPAPPPQN	9 SGDDKAEEBUNDEALEKGEAEGSEEEADEEDKAEDAKEEKYEPEKMEA 18 ESD	EDYV	BAQKLPGDPPCWTSGLPFPVPPREVIKASPHAPDPSAFSYAPPGHPLPLGLHDTARPVLP	RPPTISNPPPLISSAKHPSVLERQ	:	4 PVGPVTMGLPLPMDPKKLA	ESLG		2 TITRIIGEDSPSRLDRGREDSLPKGHVIYEGKKCHVLSYEGGMSVTQCSKEDGRSSSGPP ::	2 HETAAPKRTYDMMEGRVGRAISSASIEGLMGRAIPPERHSPHHLKEQHHIRGSITQGIPR		2 SYVEAQEDYLRREAKLLKREGTPPPPPPRDLTBAYKTQALGPLKLKPAHEG-L :	WATER TO STUDY TO BE TO TO TO TO TO TO TO TO TO TO TO TO TO	S SADDKASGRGA-ESPFEEKSGKQGSPDQVSPVSEMTSTSLYQDKQEGKSTDFADFKEDGS SADDKASGRGA-ESPFEEKSGKQGSPDQVSPVSEMTSTSLYQDKQEGKSTDFADFKEDFG		4 QEKKTDDVEAMSSQPALALDERKLGDVSPTQIDVSQFGSFKEDTKMSISEGTVSDKS
7 5	601 759	661	. 11	824 779	864	839	894	948	1008	1042	1038	1098	1145	1144	1182	1259	1242	1302	1351	1362		1415	1463	1474
~ 0	~ 0	~ ^		n .	_	٠ .							_			_			_					

8 6 6 6

8 4 8 6 8 6 8 6 8 6 8 6

6 6 6

8 4 8 4 8 4 8 4

1673 GAGVLHITENGPTEV-DYSPSDM-----QDSSLSHKIPPMEEP---SYTQDN--- 1715 1843 GGGGSSSRPASHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGIITAVEPSKPTVLRSTS 1902 1903 TSSPVRPAATFPPATHCPLGGTLDGVYPTLMEPVLLPKEA-------------- 1942 2179 1496 ARALERACYEESLKSRPGTASSSGGSIARGAPVIVPELGKPROSPLTYEDHGAPFAGHLP 1555 1531 ATPVDEGVAEDTYSHMEGVASVSTASVATSS---FPEPTTDDVSPSLHAEVGSPHSTEVD 1587 1556 RGSPVTMREPTPRLOEGSLSSSKASQDRKLTSTPREI----AKSPHSTVPEHHPHPISPY 1611 IDLSOVPHLPVLVPPTPGTPATAMDRLAYLPTAPQPFSSRHS----SSPLSPGGPTHLTK 1787 1788 PTTTS----SSERERDRDRERDRDREREKSILTSTTTVEHAPIWRPGTEOSSGSSGS 1842 1862 GKTP------GDFSYAYQKPEËTTRSPDËEDYDYESYEKTTRTSDVGGYY 1905 1943 -PRVARPERPRADIGHAF--LAKPPARSGLEPASSPSKGSEPRPLVPPVSGHATIARIPA 1999 2000 KNLAPHHASPDPPAPPASASDPHREKTQSKPFSIQELELRSLGYHGSSYSPEGVEPVSPV 2059 2060 SSPSLTHDKGLPKHLEELDKSHLEGELRPKQPGPVKLGGEAAHLPHLRPLRPESQPSSSPL 2119 2180 PSDL------YLPPPDHGAPARGSPHSEGGKRSPEPNKTSVLGGGEDGIEPVS 2226 2073 PSEARQDVDLCLVSSCEYKHPKTELSPSFINPNPLEWPASEEPTEESEKPLTGSGGAP-- 2130 2131 PPPGGKQQG--------RQCDETPPTSVSESAPSQTDSDVPPE----TE----- 2167 2168 -ECPSITADANIDSEDESE------TIPTDKTVTYKHMDPPPAPVODRSPSP 2212 2322 EHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIR 1671 1672 GYPDTAALENRQTIINDYITSQOMHHNTATAMAQRADMLRGLSPRESSLALNYAAGPRGI 1731 2285 SKKQEINKKLNTHNRNEPEYNISQPGTEIFNMPAITGTGLMTYRSQ-----AVQE---- 2334 -HASTNM-----GLEAIIRKALMGKYDQWEESPPLSANAFNPLNASASLPAAMPITAAD 2387 ----GHTLGDPSYSYETTEK----- 2024 2388 GRSDHTLTSPGGGGKAKVS---GRPSSRKAKSPAPGLASGDRPPSVSSVHSEGDCNRRTP 2444 2025 YETSTKTTRIPDISTYCYETAEKITRI-----PQ-----ASTYSYETSDLCYTAEKKS 2227 PPEGMTEPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQ--PPAFFSKLTESNSAMVK : : | | | | : | BGGYSYDISEKTTSPP-EVSGYSYEKTERSRRLLDDI---SNGYDDSB---DG-----2120 LOTAPGVKGHQRVVTLAQHISEVITQDYTRHHPQQLSAPLPAPLYSFPGASCPVLDLRRP 2445 LINRVWEDRPSSAGSTPFPYNPLIMRLQAGVMASPPPGGLPAGSGPLAGPHHA 2497 AANASASKSAKTATAGP-----GTTKTTKSSAVPPGLPVYLDLCYIPNHS 2367 1612 1732 1952 1998 g 셤 ò a ઠે a ò 셤 ò 셤 g ઠે ò ద ò g ò g ò 셤 ò g 셤 ò a g Š ò 8 8 g ò 8. 8

```
PDZ.
C2 DOMAIN 1.
C2 DOMAIN 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        788
                                                                                                                                                                                                                                                                                                                                                                                                                                    369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           518
                                         SEQUENCE
                                                                   Query Match
Best Local
  DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                       유
                                                                                                                                                                                                                                                                                                                                                                                                                                    Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
    FT
FT
SO
                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q
                                                                                                                                                                                                                                               "Aczonin, a 550-kd putative scaffolding protein of presynaptic active zones, shares homology regions with rim and bassoon and binds profilin.";
                                                                                                                                                                                                                                                                                                        J. Cell Biol. 147:151-162(1999).

-!- FUNCTION: May act as a scaffolding protein involved in the organization of synaptic active zones and in synaptic vesicle trafficking (By similarity).

-!- SUBUNIT: Interacts with Rabacl/Pral and profilin (By similarity).

-!- SUBCELLULAR LOCATION: Concentraced at the presynaptic side of
                                                                                                                      Gallus gallus (Chicken).
Gallus gallus (Chicken).
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             10 X 10 AA TANDEM APPROXIMATE REPEATS |
P-A-K-P-Q-P-C-P-X.
C4-TYPE (POTENTIAL).
POLY-PRO.
                                                                                                                                                                                                                                                                                                                                                                                 synaptic junctions (By similarity)
-!- DOMAIN: C2 domain 1 is involved in binding calcium and phospholipids. Calcium binds with low affinity but with high specificity and induces a large conformational change.
-!- SIMILARITY: Contains 2 C2 domains.
-!- SIMILARITY: Contains 1 PDZ/DHR domain.
                                                                                                                                                                                                                    MEDLINE=59439764; PubMed=10508862; MapLine B., Petrasch-Parwez B., Killmann M.W.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GO; GO:0005502; C:synaptic junction; ISS.
GO; GO:0005509; P:calcium ion binding; ISS.
GO; GO:0005544; P:calcium-dependent phospholipid binding; ISS.
GO; GO:0005522; P:profilm binding; ISS.
GO; GO:0007010; P:cytoskeleton organization and biogenesis; ISS
GO; GO:0006080; P:synaptic vesicle targeting; ISS.
InterPro; IPR000088; C2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Calcium/phospholipid-binding; Zinc; Metal-binding; Zinc-finger;
                                                                          41, Last sequence update)
41, Last annotation update)
                                   5120 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Interpro; IRR001478; PZ.
Interpro; IRR001565; Synaptotagmin.
Interpro; IRR001865; Synaptotagmin.
Interpro; IRR008899; Znf_piccolo.
Pfam; PP00168; PDZ; 1.
Pfam; PP005715; Zf_piccolo; 2.
Pfam; PP005715; Zf_piccolo; 2.
RMART; SW00228; PDZ; 1.
PROSITE; PS00049; C2 DOMAIN 1; 1.
PROSITE; PS50004; C2 DOMAIN 1; 1.
PROSITE; PS50106; PDZ; 1.
                                                                                                Piccolo protein (Aczonin) (Fragment).
                                                             28-FEB-2003 (Rel. 41, Created)
28-FEB-2003 (Rel. 41, Last seq
28-FEB-2003 (Rel. 41, Last ann
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; Y19187; CAB60725.1; -.
HSSP; P04410; 1A25.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               392
859
2343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1
357
                                       STANDARD;
                                                                                                                                                                                                          SEQUENCE FROM N.A.
TISSUE=Brain;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               368
836
2324
                                                                                                                                                                      Gallus.
NCBI_TaxID=9031;
                                                                            28-FEB-2003
28-FEB-2003
                                       CHICK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ZN_FING
ZN_FING
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Repeat.
NON TER
DOMAIN
                          PCLO_CHICK
```

```
758
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             810
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  649
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    705
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EEETAAAPP--VEEGEEOKPPAAEELAVDTGKAEEPVKSECT----EEAEEGPAKGKDA 880
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                673 MEKERNARRKKKKAPAAASEBAAFPPVVEDEEMEASGVSGNEEEMVEBAEALHASGNEVP 732
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      787
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             827
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         614 RWTEBEM-ETAKKGLLEHGRUWSAIARMVGSKTVSQCKNFYFNYKKRQNLDEILQQHKLK 672
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       567 KVIDSPMPBITKP------PADIHPAGDKPDS-------KPLPQVSRQ---- 601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----BEKDEKEKEKEAEKEEEKPEVENDKEDLL----KEKTDDTSGEDNDEKEAVAS 566
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   515
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            613
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              318 SSQTAGAAKSLAQQPGLTK-------PP-------GQQPGPEKPLQQKQA 353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MNYMSEQEKETFREKFMQHPKNFGLIASFLERKTVAECVLYYYLTKKNENYKS----- 481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  354 S---TTQPVESIPKK-----TFCPLCTTELLLH---TPEKANYNTCTQCHTV 395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   517
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       396 VCSLCGFNPNPHITEIKEWLCLNCQMQRALGGDLASGHGPGPQLPPPKQKTPTPASTAKP 455
                                                                                                                                                                                                                                                                                                                  Gaps 126;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   146 EPVSPPSPPHTDPBLELVPPRLSKEELIQNMDRVDREITMVEQQISKLKKKQQQLEBEAA 205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        266 YHENIKINQAMRKKLILLYFKRRNHARKOWKOKFCORYDQLMEALEKKVERIENNPRRR-- 323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              324 -AKESKVREYYEKQFPEIRKQRELQERMQSRVGQRGSGLS-------MSAARS 368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       258 РУКРЅАООАGРРКООРGSEKPTAOOTGPAKOPPOPGPGKTPLOOTGPVKOVPPOAGPTKP 317
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EHEVSEIIDGLSEQENLEKQMRQLAVIPPMLYDADQQRIKFINMNGLMADPMKVYKDRQV 428
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        100 MEFIESKRPRLELLPDPLLRPSPL------LATGQPAGSEDLTKDRSLTGKL 145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    51
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       706 GKLFGFGASIFSQASSLISTAGOPGSQTSGPAPPATKQPQPPSQPPASQAPP-----K
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           733 RGECSGPATVNNSSDIESIPSPHTEAAKDIGONGPKPPATLG--ADGPPGPPTPPR---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ------RTSRAPIEPT--PASEATGAPTPPPAPPSPSAPPVVPKEEK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                456 SPOLQPGQKKDASPKPDPSQQADSKKPVPQKKQPSMPGSPPVKSKQTHAEPSDTGQQIDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            567 KGRKTANSOGRRKGRITRSMANBANSBEAI-----TPQQSABLASMBL-----NESS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----LVRRSYRRR---GKSQQQQQQQQQQQQQQQQQQPMPR$SQ--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          : | : | : | : | : | : | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | : | | : | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 204 QQTSAKTSAGPTKPL----PQQPDSAKTSSQAPPTKPSLQQS--GSVKQPSQQPARQGG
                                                                                                                                                                                                                                                                                                                                                                                                      40 LEYQHHSRDYASHLSPGSIIQPQRRRPSLLSEFQPGNERSQELHLRPESHSYLPELGKSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LDTTHHPROPGKPPDPGPGLSKSRTVDVLKTEQRAPGRSPSSISLRE-----SKSR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      206 KPPEPEKPVSPPPIESKHRSLVQIIYDENRKKAEAAHRILEGLGPQVELPLYNQPSDTRQ
                                                                                                                                                                                                               Query Match
3.1%; Score 407; DB 1; Length 5120;
Best Local Similarity 18.8%; Pred. No. 6.5e-06;
Matches 577; Conservative 322; Mismatches 1162; Indels 1014;
4414 4493 PDZ.
4627 4726 C2 DOMAIN 1.
503 5094 C2 DOMAIN 2.
5120 AA, 560751 MW; A658D9891B65B412 CRC64;
```

```
KASQVEK-----TPKVESKEKVIVK----
                                                                                                                                                                                                               EMBL; U52950; AAB17068.1;
                                                                                                                                                                                                                                                                                                                                                                                                        2459
1885
1902
1919
1936
1970
2004
2021
2038
2038
786
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     127 12
140 14
2112 211
2169 216
2459 AA;
                                                                                                                                                                                                                                                                                                                                                                                    Microtubule; Repeat;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1954
1988
2005
2022
2039
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1937
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           699
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REPEAT
REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                      REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                           REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DEVELOPMENTAL STAGE, AND PHOSPHORYLATION.

MEDLINE-97405699; PubMed-9260743;

MA D., Nothias F., Boyne L.J., Fischer I.;

MA D., Nothias P., Boyne L.J., Fischer I.;

MA D., Nothias E., Boyne L.J., Fischer I.;

MA D., Nothias B., MAPLA and MAPLA B., MARLA B., MAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DEVELOPMENTAL STRAGE: In cerebral cortex, spinal cord and sciation bevelopment and are low in adults. In decrease during postnatal development and are low in adults. In dorsal root ganglia levels remain high throughout development.

INDUCTION: By nerve growth factor.

INDUCTION: By nerve growth factor with many copies of the sequence KKEE and KKEI/V, repeated but not at fixed intervals, which is responsible for the binding of MAPIB to microtubules.

FIGHT AND ADDITION IN ADDITION TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND TO SECOND T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TISSUB SPECIFICITY: Nervous system (spinal cord, brain stem, cerebellum and cerebrum). Not expressed in liver, spleen, kidney,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ç
                                     PISZOS; Q62958; Q9ER21; Q9QW92;
01-APR-1990 (Rel. 14, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
28-FBB-2003 (Rel. 41, Last annotation update)
Microtubule-associated protein 1B (MAP 1B) (Neuraxin) (Contains: MAP1 Injaht chain LCl).
                                                                                                                                                                                                                                           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PTM: Phosphorylated.
SIMILARITY: TO MAPIA.
2GUTION: A C-terminal fragment of this protein (residues 1597
2459) was originally described as neuraxin in Ref.3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TISSUE-Spinal cord;
MEDLINE-90059871; PubMed=2555150;
Riemitz A., Grenningloh G., Hermans-Borgmeyer I., Kirsch J.,
Littauer U.Z., Prior P., Gundelfinger B.D., Schmitt B., Betz H.;
"Neuraxin, a novel putative structural protein of the rat central nervous system that is immunologically related to microtubule-
                                                                                                                                                                                                                                                                                                                                                                                                                                         of the 5' end of the rat microtubule.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       domains on
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TISSUE SPECIFICITY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SECURNCE OF 96-2459 FROM N.A., DOMAIN, AND INDUCTION. STRAIN-Sprague-Dawley, TISSUE-Brain, and Glial tumor; MEDLINE-92347374; PubMed=1639092; Zaunner W., Kratz J., Staunton J., Feick P., Wiche G.; Identification of two distinct microtubule binding de recombinant rat MAP 1B.";
                                                                                                                                                                                                                                                                                                                                                                                                                     Liu D., Fischer I.; "Isolation and sequencing of the 5' end of tassociated protein (MAPIB)-encoding cDNA."; Gene 172:307-308(1996).
                                                                                                                                                                                                                                                                                                                                 [1]
SEQUENCE OF 1-142 FROM N.A.
SEQUENCE OF 1-142 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE OF 1541-2459 FROM N.A., AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Eur. J. Cell Biol. 57:66-74(1992)
                                                                                                                                                                                                                                                                                                                                                                               STRAIN=Sprague-Dawley; TISSUE=Tes;
MEDLINE=96257242; PubMed=8666295;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         associated protein 5.";
EMBO J. 8:2879-2888(1989)
                                                                                                                                                                                                                              Rattus norvegicus (Rat).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            heart or muscle.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           +++
```

```
511
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             K------PVSPPPIESKHRSLVQIIYDENRKKAEAAHRILEGLGPQVELPLYNQPSD 262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            461
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    263 TRQYHENIKINQAMRKKLILYPKRRNHARKQWKQKFCQRYDQ--LMEALEKKVERIENNP 320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LEL----VPPRLSKEELLQNMDRVDREITMVEQQISKLKKKQQQLEEEAAKPPEP----E 211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                353 İGVVFLNVPENLKNPEPNIKMKRSTEEACFTLOYLNKLSMKPEPLFRSVGNAIEPVILFQ 412
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      101 -EFIESKRPRLELLPDP-----LLRP-----SPLLATGQ-----PAGSED-- 134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 234 SESVEVPSP-PDILEPPTSGGFLKLSKPCCYIFPGGRGDSALFAVNGFNMLINGGSERKS 292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ----LTK--DRSLTG------RLEPVSPPSPPHTD-----PE 159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      293 CFWKLIRHLDRVDSILLTHIGDDNLPGINSMLQRKIAELEEESQGSTSNSDWMKNLISPD 352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RRRAKESKVREYYEKQFPEIFKQRELQERMQSRVGQRGSGLSMSAARSEHEVSEIIDGLS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     381 EQENLEKQMRQLAVIPPMLYDADQQRIKFINMNGLMADPMKVYKDR---QVMNMWSEQEKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         174 LLSTTHPANKASLTLFCPEEGDWKNSNLDRHNLQDFINIKLNSASILPEMEGLSEFTEYL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ----HLRPESHSYLPEL-GKSEM----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         .:. | .:. | .:. | .:. | .:. | .:. | .:. | .:. | .:. | .:. | .:. | .:. | .:. | .:. | .:. | .:. | .:. | .:. | .:.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CONTAINS MANY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         729;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 2459;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        M -> V (IN REF. 1).
T -> S (IN REF. 1).
R -> K (IN REF. 3).
L -> I (IN REF. 3).
W; 2E3F6872DEDB8BA2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                KKEI/V REPEATS)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BASIC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches 489; Conservative 326; Mismatches 995;
                                                                                                                                                                                                                                                                                                                                                                         MAP1 LIGHT CHAIN LC1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         No. 3.4e-06
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (HIGHLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3.1%; Score 406.5;
19.3%; Pred. No. 3.4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GLU-RICH.
LYS-RICH (
KKEE AND F
                                                                                                                                                                                                                                                                                                                                                                                            MAPIB 1.
MAPIB 2.
MAPIB 3.
MAPIB 4.
MAPIB 6.
MAPIB 6.
MAPIB 7.
MAPIB 9.
MAPIB 9.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LYS-RICH.
                                                                                                                                                                                                                                                                                                                                                      Phosphorylation.
                                                                                                                                                                                                                                                          PIR; A56577; A56577; MAPIB neuraxin.
InterPro; IPR000102; MAPIB neuraxin.
Pfam; PF00414; MAPIB neuraxin; 10.
PROSITE; PS00230; MAPIB_NEURAXIN; 8.
                                                                                                                                                                                                                    X60370; CAC16162.1; -.
X16623; CAA34620.1; ALT_SEQ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        68 LLSEFOPGNERSOEL-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      127
127
0 140
2112
2169
'; 269497 M
```

		2 SDVVPPRDM 1 SGGGGGSSS 2 TAAYQTSSS 3S 6 PGDFNYAYQ 5 VARPERPRA 6 TERTIKSPC	1988 VSGHATIAR 1960 VSGYTYEKT 2032 SIQELELRS 2015 SYETTIKTT 2074 -LEELDKSH 2075 LVSSCEFKH 2130 QRVVTLAQH	2132 QCDETPPTS 2185 LPP 2184 VTYKHMDPP 2233 2224 KALKKDLKE 2279 NSAMVKSKK 2279 NSAMVKSKK 2275 VSRVASPKK 7 28 HUMAN CA11 HUMAN CA11 HUMAN CA11 HUMAN CA11 HUMAN
		1742 1841 1893 1893 1945 1945		SULT 28
88888	888888	8 8 8 8 8 8	888888	4 6 4 6 4 6 5 4 5 4 5 4 5 4 5 4 5 4 5 4
QQQ 498 ::: EEK 652 EDN 558 : KD- 705 TTEE 618 SDT 745		EEL 904 ATT 905 VSG 956 LKQ 965 KGE 1004 APQ 1013		:: LSYE 1277 LSYE 1281 PYYQ 1330 -IEG 1329 DSES 1381 1385 FEGK 1424 LRHT 1436 :: MSSQ 1478
YLTKKUENYKSLVRRSYRRGKSQQQQQ SKPKVTKDKVVKKBIKTKPEEKKGEK KEEEKPEVENDKEDLLKEKTDDTSGEDN : : : : : KEIKKEIKKEEKKELKKEVKKETPLKD- NSERAITPQQSAELASMELNESSRWTEE : : : : : : : : : : : : : : : : : : : : : : : :	YKKRONLDEILOCHKLKMEKERN	ABSPDEGITTTGEGGGECEQTPEELEATABEGLKAEKKEGGSGRATTBYZEKAETERAEEPEBDGBDNVSGBYZEKAETERAEEPEBDGBDNVSGPAGDDRAKEDMODEALEKGEPARPPAPPPQNLQPESDAPQ	COWTSGLPFPVPPREVI 	TPLGERSYNFSLTP GHVIYEGKKGHVLA GSV-TGSAGHTP NISSAS SISPMDEPVPD REALLKREGTPP REALLKREGTPP REALLKREGTPP REALLKREGTPP TDKALGRRSESPFF ADGRSIHEIPREELI PSPEKKASDAEIM
SFLERKTVAECVLYYYLTKKNENYKSIVRRSYRRRGKSQQQQQ ::	QCKNPYPNYKRQNI 	: : : : :	APPADKEAFAAEAQKLPGDPPCWTSGLPFPVPPREVI 	TOVKTGAELDIKDVSDERLSPAKSPSLSPSPPSPIEKTPLGERSVNFSLTPNEI SITHGTPADVLYKGTITRIIGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYE
FGLIASFLERKTVAECVIVKAEVAEKAATB PMPRSSQEEKDEKEKEKE KKDEKPKEE KKDEKPKEE TANSQGRKGRITRSMANAKKEVKKDEK	NWSALARMVGSKTVSC 	VIQKET VIQKET	ADKEADKEADKE	KDVSDERLSPAKS AKGTITRIIGEDS SEQUIDANUEHC GPPHETAAP PTEUTENAQAVP EQHHIRGSITQGI EQHHIRGSITQGI LIGSEAYKTQALGPLK SDLYQDK
439 TFREKFMOHPKNFGLIASFLERKTVAECVLYYYLTKKNENYKSLVRRSYRRRGKSQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ			1014 OPG-SSPRGKSRSPAPPADKEAFABEAQKLPGDPPCWTSGLPFPVPPREVI 1014 OPG-SSPRGKSRSPAPPADKEAFABEAQKLPGDPPCWTSGLPFPVPPREVI 1064 OPGVGSPSREPASIHDETLPGGESESATASDEENREDQPEEFTATGYTQSTIEI- 1064 KASPHAPDPSAFSYAPPGHPLDLGLHDTARFVLPRPPTISNPPPLISSAKHPSVLE 1120SSEPTPMDEMSTPRDWTDETNNEETESPSQEFVNITKYESSL- 1120 RQIGAISQGMSVQLHVPYSEHAKAPVGPVTMGLPLPMDPKKLAPPSGVGDGLSPRGQ 1163YSQEYSKPVVASFNGLSDGSKTDATDGRDYNASASTISPPSSMEEDKFSKSALRD 1178 AGPPESLGVPTAQEASVLRGTALGSVPGGSITKGIPSTRVPSDS	
439 TFR 606 V-F 499 QQQ 1: 653 PKK 559 DEK		862 TEP 852 AVD 905 906 AKS 957 SAS 966 LKQ	1014 QPG 1014 QPG 1064 QPG 1120 1120 RQI 1163	1218 AYR 1222 AIT 1228 GAS 1282 GGM 1331 1332 PIE 1386
6 6 6 6 6	6 8 6 8 6 8 6	8 8 8 8 8 8	8 8 8 8 8 8 8	8 4 8 4 8 4 8 8

8	437	D 149
QQ Q	1479 SALAL	DERKLGGDGSPTQVDVSQFGSFKEDTKMSISEGTVSDKSATPVDEGA- 1530
ò	1496 ARALE	8
q	1531	EDTYSHMEGVASVSTASVAISSPPEPTTDDVSPSLHAEVGSPHSTEVD 1578
ò	9	PVTMREPTPRLOEGSLSSSKASQDRKLTSTPREIAKSPHSTVPEHHPHPISPY 1611
q	1579 DSLSV	SLSVSVVQTPTTFQETEMSPSKEECPRPMSISPPDFSPKTAKS-RTPVQDHRSE 1632
ò	1612 EHLLR	EHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIR 1671
g	1633	QSSMSIEFGQESPEHSLAMDFSRQSPDHPTV 1663
<i>8</i> €	α .	ITSQQMHHNTATAMAQRADMLRGLSPRESSLALNYAAGPRGI 173
gg	1664 GAGML	GAGWLHITENGPTEV-DYSPSDI 1685
ò		IDLSQVPHLPVLVPPTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTHLTKPTTT 1791
Q	1686 QDSS-	LSHKIPPTEEPSYTQDNDLSELISVSQVEASPSTSSAHTPSQIASPLQEDTL 1741 ·
දු දි	1792 SSSER 1742 SDVVP	SSSERERDRDRENDRDREREKSILTSTTTVEHAPIWRPG-TEQSSGSSGS 1840
ờ	-	
qq	: 1802 TAAYQ	:
ò	1893	SKPTVLRSTSTSSPVRPAATFPPATHCPLGGTLDGVYPTLMEPVLLPKEAPR 1944
qq	1856 PGDFN	PGDFNYAYQKPESTTESPDEEDYDYESHEKTIQAHDVGGYYYEK 1899
ò	1945 VARPE	VARPERPRADIGHAFLAKPPARSGLEPASSPSKGSEPRPLVPP 1987
Ор	1900 TERTI	TERTIKSPCDSGYSYBTIEKTTKTPEDGGYSCEITEKTTRTPEEGGYSYEISEKTTRTPE 1959
ò	1988 VSGHA	PPAP
qq	1960 VSGYT	SGYTYEKTERSRRİLDDISNGYDDTEDGGHTLGDCSYSYETTEKITSFPESESY 2014
ò	2032 SIQEL	SIQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDKGLPKH 2073
QQ	2015 SYETT	SYETTIKTIRSPDISAYCYETWEKITKIPQASTYSYETSDRCYİPERKSPSEARQDVDLC 2074
ολ	2074 -LEEL	-LEBLDKSHLEGELRPKOPGPVK-LGGEAAHLPHLRPLPESQPSSSPLLQTAPGVKGH 2129
Ор	2075 LVSSC	LVSSCEFKHPKTELSPSFINPNPLEWFAGEEPTEESERPLTQSGGAPPPSGGKQQGR 2131
Š	130	VITQDYTRHHPQQLSAPLPAPLYSFPGASCPV-LDLRRPPSDLY 2
q	N	IPTDKT 218
ò	2185 L	DHGAPARGSPHSEG
QQ	2184 VTYKH	VTYKHMDPPPAPWQDRSPSPRHPDVSWVDPEALAIEQNIG 2223
ò	2233	QTEPSRMGS
ф	2224 KALKK	KALKKDLKEKAKTKKPGTKTKSSSPVKKGDGKSKPSAASPKPGALKESSDK 2274
δ	2279 NSAMV	NSANVKSKKQE-INKKLNT 2296
음 :	2275 VSRVA	VSRVASPKKKESVEKAMKT 2293
RESULT CA11 HU ID CA AC PO	T 28 HUMAN CA11 HUMAN P02452; Q14037;	STANDARD; PRT; 1464 AA. ; Q15176;

```
SEQUENCE
SEQUENCE OF 1-34 FROM N.A. MEDLINE=85130970; PubMed=2857713; Chu M.-L., de Wet W.J., Bernard M.P., Ramirez F.; Chu M.-L., de Wet W.J., Bernard M.P., Ramirez F.; Fine structural analysis of the human pro-alpha 1 (I) collagen gene. Promoter structure, Alul repeats, and polymorphic transcripts."; J. Biol. Chem. 260:2315-2320(1985).
                                                                                                                                                                                                                                                                                                                                                            TISSUE=Skin;
MEDLINE=71038625; PubMed=5529814;
Click E.M., Bornstein P.;
"Isolation and characterization of the cyanogen bromide peptides from "Isolation and characterization of the cyanogen bromide peptides from the alpha 1 and alpha 2 chains of human skin collagen.";
Biochemistry 9:4699-4706(1970).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Nucleotide sequences of complementary deoxyribonucleic acids for the pro alpha 1 chain of human type I procollagen. Statistical evaluation of structures that are conserved during evolution."; Biochemistry 22:5213-5223(1983).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE=88097389; PubMed=3480516;
Bornstein P., McKay J., Morishima J.K., Devarayalu S., Gelinas R.E.
"Regulatory elements in the first intron contribute to
transcriptional control of the human alpha 1(1) collagen gene.";
Proc. Natl. Acad. Sci. U.S.A. 84:8869-8873(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE-11001508; PubMed=4319110;
Morgan P.H., Jacobs H.G., Segrest J.P., Cunningham L.W.;
"A comparative study of glycopeptides derived from selected
vertebrate collagens. A possible role of the carbohydrate in fibril
formation.":
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Maekelae J.K., Raassina M., Virta A., Vuorio E.;
"Human pro alpha 1(I) collagen: cDNA sequence for the C-propeptide
                                                                                                                                                                                                                                 SEQUENCE OF 1-181 FROM N.A. MEDLINE=84270697; PubMed=6462220; Chu M.-L., de Wet W.J., Bernard M.P., Ding J.-F., Morabito M., Myers J., Williams C., Ramirez F.; "Human pro alpha 1(1) collagen gene structure reveals evolutionary when pro alpha 1(2) collagen gene structure reveals evolutionary conservation of a pattern of introns and exons.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEGUENCE OF 425-1464 FROM N.A.
MEDLINE-84080385; PubMed-6689127;
Bernard M.P., Chu M.-L., Myers J.C., Ramirez F., Eikenberry E.F.,
                                                                    Homo sapiens (Human).
Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                              SEQUENCE OF 1-472 FROM N.A. MEDLINE=89025644; PubMed=3178743; Tromp G., Kuivaniemi H., Stacey A., Shikata H., Baldwin C.T., Jaeniach R., Prockup D.J.; Structure of a full-length cDNA clone for the prepro alpha 1(I) chain of human type I procollagen."; Biochem. J. 253:919-922(1988).
 21-JUL-1986 (Rel. 01, Created)
01-MAR-1989 (Rel. 10, Last sequence update)
10-OCT-2003 (Rel. 42, Last annotation update)
Collagen alpha 1(I) chain precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Nucleic Acids Res. 16:349-349(1988)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Biol. Chem. 245:5042-5048(1970)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE=88124208; PubMed=3340531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE OF 1229-1454 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE OF 1-34 FROM N.A.
                                                                                                                                                                                                                                                                                                                      Nature 310:337-340(1984).
                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE OF 263-268.
TISSUE=Skin;
                                                                                                                                                                                                                                                                                                                                                SEQUENCE OF 162-301.
                                                                                                           NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Prockop D.J.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         rissum=Bone
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2
```

```
MEDLINE=88033031; PubMed=3667599; MEDLINE=88033031; PubMed=3667599; Vogel B.E., Minor R.R., Freund M., Prockop D.J.; Vogel B.E., Minor R.R., Freund M., Prockop Generits glycine 748 "A point mutation in a type I procollagen gene converts glycine 748 of the alpha 1 chain to cysteine and destabilizes the triple helix in a lethal variant of osteogenesis imperfecta.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VARIANT OI-II ARG-569.
MEDLINE-8722295; PubMed=3108247;
Bateman J.F., Chan D., Walkers I.D., Rogers J.G., Cole W.G.;
Bateman J.F., Chan Do: walkers imperfects due to the substitution of arginine for glycine at residue 391 of the alpha 1(I) chain of type i
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE=97255959; PubMed=9101290; Kuivaniemi H., Tromp G., Prockop D.J.; Kuivaniemi H., Tromp G., Prockop D.J.; Kuivaniemi H., Tromp G., Prockop D.J.; Mutations in fibrillar collagens (types I, II, III, and XI), fibrillassosofiated collagen (type IX), and network-forming collagen (type X) cause a spectrum of diseases of bone, cartilage, and blood vessels."; Hum. Mutat. 9:300-315(1997).
                                                                                                                                                                                                                                                                                                                                                               MEDINE-91184577; PubMed=2010058; Kuivaniemi H., Tromp G., Prockop D.J.; "Mutations in collagen genes: causes of rare and some common diseases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VARIANT OI-II ARG-842.
MEDLINE=88298828; PubMed=3403550;
Bateman J.F., Lamande S.R., Dahl H.H., Chan D., Cole W.G.;
"Substitution of arginine for glycine 664 in the collagen alpha 1(I)
"Substitution of perinatal osteogenesis imperfecta. Demonstration of the peptide defect by in vitro expression of the mutant cDNA.";
J. Biol. Chem. 263:11627-11630(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDILINE 89218628; PubMed=3244312; Labhard M.E., Wirtz M.K., Pope F.M., Nicholls A.C., Hollister D.W.; Labhard M.E., Wirtz M.K., Pope F.M., Nicholls A.C., Hollister D.W.; Aysteine for glycine substitution at position 1017 in an alpha [1 (1) chain of type I collagen in a patient with mild dominantly inherited osteogenesis imperfecta.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE-8287390; PubMed-3016737; Cohn D.H., Byers P.H., Steinmann B., Gelinas R.E.; Cohn D.H., Byers P.H., Steinmann B., Gelinas R.E.; Lethal osteogenesis imperfecta resulting from a single nucleotide change in one human pro alpha 1(1) collagen allele."; Proc. Natl. Acad. Sci. U.S.A. 83:6045-6047(1986).
                                  MEDLINE-88033098; PubMed-2822714;
Rossouw C.M.S., Vergeer W.P., du Plooy S.J., Bernard M.P., Ramirez
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE=91374476; PubMed=1895312;
Byers P.H., Wallis G.A., Willing M.C.;
"Osteogenesis imperfecta: translation of mutation to phenotype.";
J. Med. Genet. 28:433-442(1991).
                                                                                                                                             "DNA sequences in the first intron of the human pro-alpha 1(I)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            "The human type I collagen mutation database.";
Nucleic Acids Res. 25:181-187(1997).
                                                                                                                                                                                                collagen gene enhance transcription."; J. Biol. Chem. 262:15151-15157(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  J. Biol. Chem. 262:7021-7027(1987)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REVIEW ON OI VARIANTS.
MEDLINE=97169389; PubMed=9016532;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FASEB J. 5:2052-2060(1991).
OF 1-44 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VARIANT OI-II CYS-1166.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REVIEW ON OI VARIANTS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VARIANT OI CYS-1195.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REVIEW ON VARIANTS.
                                                                                                                                                                                                                                                                                                                                 REVIEW ON VARIANTS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Dalgleish R.;
                                                                                                                        de Wet W.J.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        in humans.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 collagen.
```

```
615 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1016
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1630
 918
                                                                                                                1030
                                                                                                                                                                                                 433
                                                                                                                                                                                                                                                         471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  838
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          871
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                913
                                                                                                                                         g
                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                      셤
                                                      ઠ
                                                                            요
                                                                                                              ð
                                                                                                                                                                      ઠે
                                                                                                                                                                                            a
                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                          名
                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8.8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                  Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                857
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          917
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AGKPGR-----PGERGPPGPQGARGLPGTA--GLPGMKGHRGFSGLDGAKGDAGP---- 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EALHASGNEVPRGECSGPATVNNSSDTESIPSPHTEAAK-DTGQNGPKPPA-TLGADGPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GQPGLPGPPGPP------GPPGPPGLGGNFAPQLSYGYDEKSTGGISVP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                811 APPSPSAP----PPVVPKEEKEETAAAPPVEEGE-----EQKPPAAEELAVDTGK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSSGAPQDSDS
                                                                                                                                       MEDLINE=89308591; PubMed=2745420;
Marini J.C., Grange D.K., Gottesman G.S., Lewis M.B., Koeplin D.A.;
"Osteogenesis imperfecta type IV. Detection of a point mutation in
one alpha 1(1) collagen allele (COLIA1) by RNA/RNA hybrid analysis.";
J. Biol. Chem. 264:11893-11900(1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GPMGPSGPRGLPGPPGAPGPQGFQ----GPPGEPGEPGASGPMGPRGPPGPPGRNGDDGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             by
of
                                                                                                                                                                                                                                                                                                                                                   Pack M., Constantinou C.D., Kalia K., Nielsen K.B., Prockop D.J.; "Substitution of serine for alpha 1(1)-glycine 844 in a severe variant of osteogeneeis imperfecta minimally destabilizes the triple helix of type I procollagen. The effects of glycine substitutions on thermal stability are either position of amino acid specific."; J. Biol. Chem. 264:19694-19699(1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                   Patterson E., Smiley E., Bonadio J., "RNA sequence analysis of a perinatal lethal osteogenesis imperfecta
                                                                                                                                                                                                                                                                                                                                                                                                                                                       VARÍANT OI-II CYS-1082.
MEDLINE-89109573; PubMed=2913053;
CODSTEAMTIOU C.D., Nielsen K.B., Prockop D.J.;
"A lethal variant of osteogenesis imperfecta has a single base mutation that substitutes cysteine for glycine 904 of the alpha 1(I) chain of type I procollagen. The asymptomatic mother has an unidentified mutation producing an overmodified and unstable type I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        "Osteogenesis imperfecta. The position of substitution for glycine cysteine in the triple helical domain of the pro alpha 1(I) chains type I collagen determines the clinical phenotype.";
                                                                                                                                                                                                                                       MEDLINE=89380165; PubMed=2777764; Lamande S.R., Dahl H.-H.M., Cole W.G., Bateman J.F.; Medracterization of point mutations in the collagen COLIA1 and COLIA2 genes causing lethal perinatal osteogenesis imperfecta."; J. Biol. Chem. 264:15809-15812(1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE=90009313; PubMed=2794057;
Starman B.J., Eyre D., Charbonneau H., Harrylock M., Weis M.A.,
Weiss L., Graham J.M., Byers P.H.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       647;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DB 1; Length 1464;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                           VARIANTS OI-II ALA-1106; VAL-1151; ARG-1154 AND VAL-1184.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          780 -----PGPPTPPRRTSRAPIEPTPASEATGAPTPP------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3.0%; Score 400.5; DB 1;
21.5%; Pred. No. 3.3e-06;
1ve 102; Mismatches 642;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VARIANTS OI CYS-272; CYS-704 AND CYS-896.
                                                                                              Biol. Chem. 264:10083-10087(1989)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Clin. Invest. 84:1206-1214(1989)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Clin. Invest. 83:574-584(1989).
                                       MEDLINE-89255493; PubMed=2470760;
                                                                                                                                                                                                                                                                                                                                          MEDLINE=90062068; PubMed=2511192;
 Biol. Med. 5:197-207(1988)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       21.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      380; Conservative
                                                                                                                          VARIANT OI-IV SER-1010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              VARIANT OI-II CYS-422
                                                                                                                                                                                                                                                                                                                           VARIANT OI SER-1022
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           procollagen.";
                                                                                 mutation."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  722
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          828
셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      કે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ઠે
```

1076 1077 YAPPGHPLPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHVP 1136 1573 1738 1851 1294 GRSSSGPPHETAAPKRTYDMMEGRVGRAISSASIEGLMGRAIPP-ERHSPHHLKEQHHIR 1352 1411 HEGLVATVKEAGRSIHEIPREELRHTPELPLAPRPLKEGSITQGTP----LKYDTGASTT 1466 1912 TFPPATHCPLGGTLDGVYPTLMEPVLLPKBAPRVARPERPRADTGH--------- 1957 SATCSADEVDEAEGGDKNRLLSP----RPSLLTPTGDPRANASPQKPLDLKQLKQRAAAI 973 560 GQDGRPGPPGPPGARGQAGVMGFPGPKGAAGEPGKAG----ERGVPGPPGAVGPAGKDG 1353 GSITQGIPRSYVEAQEDYLRREAKLLKREGTPPPPPPPSRDLTEAYKTQALGPLKLK--PA 1467 GSKKHD-----VRSLIGSPGR-----TFP--PVHPLDVMADARALERACYEESLKSRPG ----AGPKGEPGSPGENGAPGQMGPRGLPGERGRPGAP-----GPAGAR 974 PPIQVTKVHEPPREDAAPTKPA-PPAPPPQONLQPESDAPQQPGSS--PRGKSRSPAPP-GNDGATGAAGPP----GPTGPAGPPGFPGAVGAKGEA-GPQGPRGSEGPOGVRGEPGPPG ADKEAFAAEAQKLPGDPPCWTSGLP-FPVPPREVIKASPHAPDPSAFS 1137 YSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEASVLR ----GPPGERGGPGSRGFPGADGVAGPK 1197 GTA--LGSVPGGSITKGIP-STRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPS 506 GPAGERGS-PGPAGPKGSPGEAGRPGEAGLPGAKGLT-GSPGSPGPDGK----TGPPGPA 1254 RLD-----SGREDSL----PKGHVIYEGKKGHVLSYEGGM----SVTOCSKED ----QGVPGDLGAPGPSGARGERGFPGERGVQGPPGPA------GPRGANGAPG 705 NDGAKG---DAG----DAG----APGAP-----GS--QGAPGERGAAGL 1514 TASSSGGSIARGAPVIVPELGKPROSPLTYEDHGAPFAGHLPRGSPVTMREPTPRLQEGS 1574 LSSSKASQDRKLTSTPREIAKSPHSTVPEHHPHPI----SPYEHLLRGVSGVDLYRSHIP 1690 ITSQOMHNTATAMAQRADML------RGLSPRESSLALNYAAGPRGIIDLSQVP 1739 HLPVLVPPT--PGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTHLTKPTTTSSSER 1852 ASHSHAHQHSPISPRTQDALQQRPSVLHNTGWKGIITAVEPSKPTVLRSTSTSSPVRPAA -----AGPVGPAG 378 PAGAAGPAGNPGADGQPGAKGANGAPG----IAGAPGFPGARGPSGPQGPGGPPGPKGN -----GPAGEEGKRG-----------KGESG 786 PSGPAGPTGARGAPGDRGEPGPP--GPAGFA--GPPGAD----GOPGAKGEPGDAGAKGD -----AGPPGPAGPPGPIGNVGAPGAKGARGSAGP LAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAALENRQTIINDY ------TGPAGRPGEVGPPGPAGEKGSPGADGPAGAPGTPGPQGIAGQRGVV GLPGQRGERGFPGLPGPSGE----PGKQGPSGERGPPGPMGPPGLAGPPGESGRE-1797 ERDRDREREDRDREREKSILTSTTTVEHAPIWRPGTEQSSGSSGSSGGG-----GSSSRP ----GAPGAEGSPGRDGSPGAKGDRGETGPAGPP ------pgatgfpgaagrvgp----pgps-gnagpp----gppggagkeggkgprge-EAGAQGPP-GPAGPAGE----RGEQGPA-GSPGFQGLPGPAGPPGEAGKPGE------AGKSGDRGETGPAGP-739 PGPKGDRGDAGPKGADGSPGKDGVRGLTGPIGPPGPAGAPGD-SGEPGAPGSKG--DTGAKGEPGPVGVQGPP-----ARGEPGPT--GLP-----GAPGAPGAPGPVGP-

```
ANK repeat; REPEAT
                                                                                                    Name=3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REPEAT
----TC--R 1260
                                                                                                                                                                                                                                                DLRRPPSDLYLPPPDHGAPARGSPHSEGGKRSPEPNKTSVLGG------GEDGIEPV 2225
                                                                                                                                                                                                                                                                                                2226 SPPEGM-----TEPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQPPAFFSKLTESN 2279
                                                                                                                                                                                                                                                                                                                        1304 OPSVAQKAWYISKNPKDKRHVWFGESMTDGFQFEYG-----GQGSDPADVAIQLT--- 1353
                                      -AFLAKPPARSGLEPASSPSKGSEPRPLVPPVSGHATIARTPAKNLAPHHASPDPPAPPA 2016
                                                                                                                                                                                           -------RVVTLAQHISEVITQDYTRHHPQQLSAPLPAPLYSFPGASCPVL 2174
                                                                                                                                                                                                                                                                       1261 DLKMCHSDW------KSGEYWIDPNQGCNLDAIKVFCNMETGETCVYPT 1303
             -----AGPOGPRGDKGETGEQGDRGIKGHRG 1112
                                                                                        SASDPHREKTQSKPFSIQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDKGLPKHLEE 2076
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sheng M., Kim E.;
"The Shank family of scaffold proteins.";
J. Cell Sci. 113:1851-1856(200)
J. Cell Sci. 113:1851-1856 an adapter protein in the postsynaptic
-!- FUNCTION: Seems to be an adapter protein in the postsynaptic
density (PSD) of excitatory synapses that interconnects receptors
of the postsynaptic membrane including NMDA-type and metabotropic
glutamate receptors via complexes with GKAP/PSD-95 and Homer,
respectively, and the actin-based cytoskeleton. May play a role in
the structural and functional organization of the dendritic spine
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    and synaptic junction.
SUBUNIT: May homomultimerize via its SAM domain (By similarity).
Interacts with SSTR2 C-terminus via the PDZ domain. Interacts with SPTAN1, Homer-1 and DLGAP1/GKAP isoforms 1 and 2 (By similarity).
Is part of a complex with DLG4/PSD-95 and DLGAP1/GKAP (By similarity).
                                                                                                                                                                                                                                                                                                                                                                                                                              SHAI HUMAN STANDARD; PRT; 2161 AA.
09Y566; Q9NYW9;
28-FEB-2003 (Rel. 41, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
SH3 and multiple ankyrin repeat domains protein 1 (Shank1)
(Somatostatin receptor interacting protein) (SSTRIP).
                                                                                                                                           2077 LDKSHLEGELRPKQPGPVKLGGEAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQ-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3), AND INTERACTION WITH SSTR2. TISSUE-Fetal brain, Hippocampus, and Thalamus;
MEDLINE=20020275; PubMed=10551867;
Zitzer H., Hoenck H.-H., Bacchner D., Richter D., Kreienkamp H.-J.;
"Somatostatin receptor interacting protein defines a novel family of multidomain proteins present in human and rodent brain.";
J. Biol. Chem. 274:32997-33001(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               .... oupta.m. numan...
Eukaryota, Metavoma...
Kararyota, Eutheria, Frimates, Catarrhini, Hominidae, Homo.
                                                                                                                  1146 SAGAPGKDGLNGLPGPI-----GPPGPRGRTGDAGPVGPPGPPG-----
                                                                                                                                                                                                                      1218 ADDANVVRDRDLEVDTTLKSLSQQIENIRSPEGSRKNPAR----
                                                                                                                                                                                                                                                                                                                                                   2280 SAMVKSKKQEINKKLNTHNRNEPEYNISQPG 2310
                                                                                                                                                                                                                                                                                                                                                                     :: | :: | | | | 1354 --FLRLMSTEASQNITYHCKNSVAYMDQQTG 1382
                                                                 FSGLOGPPGPPGSPGEQGPSGASGP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE=20267867; PubMed=10806096;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=9606;
                1083 ARGP
                                                                                           2017
                                                                                                                                                                                                                                                2175
                                                                                                                                                                                               2131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REVIEW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SHANKI
                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                      ð
                                                               В
                                                                                           à
                                                                                                            g
                                                                                                                                            8
                                                                                                                                                                  8
                                                                                                                                                                                     ☆
                                        à
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Missing (in isoform 2).
/FITd=VSP 006069.
/FITd=VSP 006069.
-> MOLMALEGORGGGLEGGQPLCLMMSSPLPPPPFFSC
LPA (in isoform 2).
/FITd=VSP_006070.
                                                                                                                                                                                                                                        1807d=Q9Y566-3; Sequence=VSP_006071;
-!- TISSUE SPECIFICITY: Expressed in brain particularly in the amyddala, hispocampus, substantia nigra and thalamus. Isoform seems to be expressed ubiquitously.
-!- SIMILARITY: BELONGS TO THE SHANK FAMILY.
-!- SIMILARITY: Contains 6 ANK repeats.
-!- SIMILARITY: Contains 1 PDZ/DHR domain.
-!- SIMILARITY: Contains 1 sterile alpha motif (SAM) domain.
-!- SIMILARITY: Contains 1 sterile alpha motif.
SUBCELLULAR LOCATION: Cytoplasmic; postsynaptic density
                                                                                                                                                                 Name=2; Synonyms=B;
IsoId=Q9Y566-2; Sequence=VSP_006069, VSP_006070;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Alternative splicing.
                         neuronal cells (By similarity).
ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=3;
                                                                                                                                     IsoId=Q9Y566-1; Sequence=Displayed;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ANK 1.

ANK 2.

ANK 3.

ANK 4.

ANK 5.

ANK 5.

ANK 6.

SH3.

PDZ.

PDZ.

PDZ.

PDZ.

PDZ.

POLY-HIS.

POLY-GLY.

POLY-GLY.

POLY-GLY.

POLY-GLY.

POLY-GLY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Prodom; P000066; SH3; 1.
SYART; SW00228; ANK; 6.
SYART; SW00228; PDZ; 1.
SWART; SW00454; SAM; 1.
SWART; SW00356; SH3; 1.
PROSITE; PS50088; ANK REPEAT; 3.
PROSITE; PS50297; ANK_REP_REGION; 1.
PROSITE; PS50105; SAM_DOMAIN; 1.
PROSITE; PS50105; SAM_DOMAIN; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; AF163302; AAD45121.1; -. EMBL; AF226728; AAF35887.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Genew; HGNC:15474; SHANK1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR02110; ANK.
InterPro; IPR001478; PDZ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Interpro; IPR001478; PDZ
Interpro; IPR001460; SAM
Interpro; IPR001452; SH3
Pfam; PF00023; ank; 7.
Pfam; PF00595; PDZ; 1.
Pfam; PF00536; SAM; 1.
Pfam; PF00188; SH3; 1.
                                                                                                              Name=1; Synonyms=A;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           654
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HSSP; P06241; 1SHF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MIM; 604999; -
```

1216	UY 1416 A : Db 1259 S'	Qy 1470 KI	31.01		Ø, 0/51 Ø	: Db 1418 R	Qy 1619 S(1474	1658	1530	1696	1590		1645	1796	UD 1676 E		1 00	QY 1909 PA	077	1806	2022	1864	2081	1923	Qy 2137 QI	Db 1969 T	- 2195 -	Db 2016 R	oy ,** 2230 -	Db 2063 G	RESULT 30
VARSPLIC 646 654 Missing (in isoform 3). SEQUENCE 2161 AA; 225019 MW; 5PEFC ⁹ 69CBE98701 CRC64;	Query Match 3.0%; Score 399; DB 1; Length 2161; Best Local Similarity 20.7%; Pred. No. 5.2e-06; Matches 421; Conservative 191; Mismatches 775; Indels 643; Gaps 101;	RGKSQQQQQQQQQQQQQQQPPRSSQEEKDEKEREREBERPBVBNDKEDLLKEK	469 QGQSQPSAPTTKLSSGTLRSASSPRGARARSPSRGRHPEDAKRQPRGRPSSSGTPREG 526	TDDTSGEDNDEKEAVASKGRKTANSQGRKKGRITRSMAN-EANSEEAITPQQSA	FAGGIGGOCGERGODEGONGNICEKRALIDAV FGAGNEMAVADIÇAÇEGELSEDNGE	603 ELASMELNESSKYTEEEMETAKKULLEHGKNWSALAKMVGSKTVSQCKNFYFNYKKKQNL 662 1:	DEILQQHKLKMEKERNARRKKK-KAPAAASEEAAFPPVVEDEEMEASGVSGNEEEMVEEA	EEVANRSQESKQESRSDKAKRLPRHYTVGSYDSFDAPSLMDGIGPGSDYIIKEK	722 EALHASGNEVPRGECSGPATVNNSSDTESIPS-PHTEAAKD 761		762TGQNGPKPPATLGADGPPPGPPTPPRRTSRAPIEPTPASEATGAPTPPPAPPSPSA 817		818 PPPVVPKEEKEEETAAAPPVEEGEEGKPPAAEELAVDTGKAEEPVKSECTEEAEEGPAKG 877	778 PPPTISLRSKSMTSELEEMBYEQQPAPVPSMEKKRTVYQMALN 820	878 KDAEAAEATAEGALKAEKKEGGSGRATTAKSSGAPQDSDSSATCSADEVDEAEGGDKNRL 937	821 KLDEIL-AAAQQTISASESPGPGGLASLGKH	938 LSPRPSLLTPTG-DPRANASPQKPLDLKQLKQRAAAIPPIQVTKVHBPPRBDAAPTKP 994	851 -RPKGFFATESSFDPHHRAQPSYERPSFLPPGPGLMLRQKSIGAAEDDRPYL 901	995 APPAPPPPQNLQPESDAPQQPGSSPRGKSRSPAPPAD 1031	902 APPAMKESRSLSVPGSEDIPPPPTTYPPEPPYSTPTVPSSSGRLTPSPRGGPFNP 956	1032 KEAFAAEAQKLPGDPPCWTSGLPPPVPPREVIKASPHA-PDPSAFSYAPPGH 1082	957GSGGPLPASSPATPDGPSPPDTRVGSREKSLYHSGPLPPAHHPPHHHH 1006	1083PLPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLH 1134	1007 НАРРЮРНИННАНРРНРРЕМЕТССSРDDРРРКLALGPQPSLRСWRCGGPSPTPG 1060		APSPSHHGSAGGGGSSQGPALRYIQLPPRAASAAMYVPARSGRGRKGP 110	TAQBASVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYKGTITR	LVKQTKVEGEPQKGGGLPPAPSPSSPASPS	IIGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETA : : : :	QPPPAVAAPSEKNSIPIIIKAPSTSSSGRSSQGSSTEAE:	RHSPHHLKEQHHIRGSITQGIPRSYVE 136	AGEDYLRREAKLLKREGTPPPPPPSRDLTEAYKTQALGPLKLKPAHEG
FT SO	Omz	ò	đ	8 8	3 8	S 8	ò	a	ઠે	a	∂	g	ò	qq	ò	q	ò	g	ò	ДQ	ò	đ	ò	d d	ò	පු	Š	요.	ò	a	8	8 8

ATVK-----EAGRSIHEIPREELRHTPELPLAPRPLKEGSITQGTPLKYDTGASTTGSK 1469 NPTYPH-----LYPP----YLIRGYPDTAALENRQTIINDYITSQQM 1695 OVPHIPVLVPPTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTHLTKPTTTSSSE 1795 --PHPP--GPPAPAPA-----PAAPQP------GP---DPPPGTDSGI 1675 OS-KPPSIQELELRSLGYHGSSYSPEGVEPVSPVSS--PSLTHDKG---LPKHLBELDKS 2080 ISPSASSSSTSTR-HLQGVEFEMRPPL------LRRAPSPSLLPASEHKVSPAP 2015 :| : | : | | : | | STDAGDEDGGDGGLGTGAAPGPRLRHS------KSIDEGMFSAEPYLRLESAGSGAGYG 1311 KHDV--RSLIGSPGRT----FPP---VHPLDVM----ADARALERACYEESLKSRPGTA 1515 SSSGGSIARGAPVIVPELGKPR-QSPLTYEDHGAPFAGHLP----RGSPVTMREPTPRL 1569 ZEGSLSSSKASQDRKLTSTPREIAKSPHSTVPEHHP------HPISPYEHLLRGV 1618 SGVDLYRSHIPLAFDPTSIP-----RHLAP 1657 -----AMAQRADMLRGLSPRESSLALNYAAGPRGIIDLS 1735 HPLPDTPAPATPLPPVPPPAVAAAPPTLDSTASSLTSYDSEVATLTQGASAAPGD---- 1644 RERDRORERDROREREKSILTSTTTVEHAPIWRPGTEQSSGSSGGGGGSSSRPASHS 1855 AAHQHSPISPRTQDALQQRPSVLHNTGMKGIITAVEPSKPTVL-----RSTSTSSPVR 1908 : | : | : | : | 1985 | 1989 | 1985 | 1989 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | 1986 | PAATFPPATHCPLGGTLDGVYPTLMEPVLLPKEAPRVARPERPRADTGHAFLAKPPARSG 1968 LEPASSPSKGSEPRPLVPPVSGHATIARTPAKNLAPHHASPDPPAP--PASASDPHREKT 2026 PPTAGVAGGPVAVEPEVPPVP-LPTASCLPRK-LLPWEEGPGPPPPLPGPLAQPQASAL 1863 HL----EGELRPKOPGPVKLGGEAAHLPHLRPLPESQPSSSPLLQTAPGVKGHORVVTLA 2136 OHISEVITODYTRHHPQQLSAPLPAPLXSFPGASCPVLDLRRPPSDLYLPPPDHG-APA- 2194 -----RGSPHSEGGKRSPENKTSVLGGGEDGIEPV-SPPE--- 2229 RPSSLPILPSGPLYPGLFDIRGSP------TGGAGGSADPFAPVFVPPHPGIS 2062 ---GMTEPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNT--SQPPAFFSK 2274 HHNTAT------

```
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
RN 131

SEQUENCE OF 1-613 AND 4803-4903 FROM N.A.

SERUNA-ES7BL/60; TISSUE=Embryo;

RA MELINE-Z5784683; DubMed=12466851;

RA MELINE-Z5784683; DubMed=12466851;

RA MELINE-Z578683; DubMed=12466851;

RA MARAIO I., OBATON N., SAITOR R., SUZNKI H., YAMMANAKA I., KIYOSAWA H.,

RA MARAIO I., OBATON N., SAITOR R., SUZNKI H., YAMMANAKA I., KIYOSAWA H.,

RA BAIR E., Dragani T.A., MATENDA C., Corbani L.E., COUSINS S.,

RA BAIR E., Dragani T.A., Fletcher C.F., FORTEST A., Frazer K.S.,

RA Grimmond S., Gurincich S., Hirokawa N., Jackson I.E., COUSINS S.,

RA Grimmond S., Gurincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,

RA MARAIA T., Kawasawa Y., Kedzierski R.M., King B.L.,

RA MARAIA T., MARAIAS L., MARCHIONNI L., MCROZIC B.L., MIKH H.,

RA MARAIA T., NUMBAR X., ORIGHOUS J., REND S. D.,

RA RAVASI T., Reed J.C., Reed D.J., Reid J., Ring B.L., Miki H.,

RA RAVASI T., Reed J.C., Reed D.J., Reid J., Ring B.L., Miki H.,

RA RAVASI T., Reed J.C., Reed D.J., Reid J., Ring B.L., Miki H.,

RA RAVASI T., Reed J.C., Reed D.J., Reid J., Ring B.C., Ringwald M.,

RA RAVASI T., Reed J.C., Wang Y., Watanabe Y., Wells C.,

RA MARAIA S., Schneider C., Semple C.A., Setou M., Shimada K.,

RA MILANA R., Yamushaw-Boris A., Yanagisawa M., Yang I.,

WILL R., Walki K., Xawai J., Aizawa K., Azakawa T., Pukuda S.,

HILING L.G., Wynshaw-Boris A., Yanagisawa M., Yang I.,

Watazaki A., Bakai K., Sasaki.D., Shibata K., Ainayala I.,

MARAINSHI Y., Sakai K., Sasaki.D., Shibata K., Ainayala M.,

A Marazaki G. H., Waterston R., Lander E.S., Rogers J.,

RA MARAZAKI Y., Max M., Waterston R., Lander E.S., Rogers J.,

RA MARAZAKI Y., Rawai J., Aizawa K., Azakawa T.,

RA MARAZAKI Y., Sasaki.D., Shibata K., Ainayaya A.,

MARAZAKI Y., Water L., Walki W., Materston R., Lander E.S., Rogers J.,

RA MARAZAKI Y., Rawai J., Aizawa K., Azakawa T., Pukuda S.,

RA MARAZAKI Y., Waterston R., Lander E.S., Rogers J.,

RA MARAZAKI A., Maxaliyaki Y.,

RA MARAZAKI A., Maxaliyaki Y.,

RA MARAZAKI A., Watanaba Y., Walliyaki Y.,

RA MARAZAKI A., Watanaba Y., Walliyaki Y., Max
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           S-adenosyl-L-homocysteine + histone N(6)-methyl-L-lysine.
SUBUNIT: Belongs to the ASC-2/NCOA6 complex (ASCOM), which
contains ASC-2/NCOA6, the retinoblastoma-binding protein RBQ-3/
RBBPS, alpha- and beta-tubulins, the trithorax group proteins MLL2
and MLL3, and ASH2/ASCL2. Interacts with histone H3 (By
              MLL3 MOUSE STANDARD; PRT; 4903 AA.
QBBRH4, QBBK12; QBC6M3; Q923H5; Q923H6;
10-OCT-2003 (Rel. 42, Last sequence update)
15-MAR-2004 (Rel. 43, Last amoutation update)
Myeloid/lymphoid or mixed-lineage leukemia protein 3 homolog (Histonelysine N-methyltransferase, H3 lysine-4 specific MLL3) (EC 2.1.1.43).
                                                                                                                                                                                                                                                                                                                        Tan Y.C., Chow V.T.;

"Novel human HALR (WLL3) gene encodes a protein homologous to ALR and to ALL-1 involved in leukemia, and maps to chromosome 7q36 associated with leukemia and developmental defects.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CATALYTIC ACTIVITY: S-adenosyl-L-methionine + histone L-lysine =
                                                                                                                                                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. NCBI_TaxID=10090;
                                                                                                                                                                                                                    CONCEPTUAL TRANSLATION OF 814-956 AND 1377-4213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SUBCELLULAR LOCATION: Nuclear (Probable)
                                                                                                                                                                                                                                                                                   SEQUENCE OF 957-1376 AND 4214-4894 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                           Cancer Detect. Prev. 25:454-469(2001).
                                                                                                                                                                                                                                                     Unpublished observations (SEP-2003)
                                                                                                                                                                                                                                                                                                     TISSUE=Myeloma;
MEDLINE=21574953; PubMed=11718452;
                                                                                                                                             Mus musculus (Mouse)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          similarity)
                                                                                                                                                                                                                                         Blatter M.
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
The SET domain interacts with histone H3 but not H2A, H2B and may have a H3 lysine specific methylation activity (By
                                                              -!-SIMILARITY: Belongs to the transcription factor trithorax family.
-!-SIMILARITY: Contains 1 DHHC-type zinc finger.
-!-SIMILARITY: Contains 6 PHD-type zinc fingers.
-!-SIMILARITY: Contains 1 Post-SET domain.
-!-SIMILARITY: Contains 1 RING-type zinc finger.
-!-SIMILARITY: Contains 1 RING-type zinc finger.
-!-SIMILARITY: Contains 1 FRING-type zinc finger.
-!-CAUTION: The regions from 814 to 956 and 1377 to 4213 were deduced from the genomic sequence by similarity to the human
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            rransferase; Methyltransferase; Chromatin regulator; Activator; DNA-binding; Nuclear protein; Transcription regulation; Coiled coil;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -> VSDFLICF (IN REF. 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        S -> C (IN REF. 3; BAC35712).
MISSING (IN REF. 2; AAK70214).
W; 0B896490B081BA6C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COLLED COLL (POTENTIAL).
COLLED COLL (POTENTIAL).
COLLED COLL (POTENTIAL).
COLLED COLL (POTENTIAL).
COLLED COLL (POTENTIAL).
COLLED COLL (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DHHC-TYPE.
PHD-TYPE 3.
PHD-TYPE 4.
PHD-TYPE 5.
PHD-TYPE 6.
A.T HOOK (BY S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RING-TYPE.
PHD-TYPE 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PHD-TYPE 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ASP-RICH.
GLN-RICH.
NCRICIEC -
BAC32109).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CYS-RICH.
CYS-RICH.
PRO-RICH.
GLN-RICH.
PRO-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POST-SET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ENGRIT; SM00249; PHD; 3.

PROSITE; PS00354; HMGI Y; 1.

PROSITE; PS50868; POSITE; PSCOSITE; PS50280; SET; 1.

PROSITE; PS50216; ZF DHHC; 1.

PROSITE; PS01359; ZE PHD 1; 5.

PROSITE; PS50016; ZF PHD 2; 6.

PROSITE; PS50089; ZF PHD 2; 6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AA; 540182 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MGD; MGI:2444959; Mll3.
InterPro; IRR000637; AT hook.
InterPro; IRR001965; Znf PHD.
InterPro; IRR001841; Znf ring.
Pfam; PF02178; AT hook; I.
Pfam; PF00628; PHD; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                     3193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Repeat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              675
4713
4903
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1831
2682
3022
                                                   similarity).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1330
1743
3047
3166
3224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 970
                                                                                                                                                                                                                                                     sednence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Zinc-finger;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ZN_FING
ZN_FING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ZN_FING
ZN_FING
ZN_FING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
```

2313 нг					1014				1095 VI 2590 DE	1135 VE	2647 GE	1193 87	2692 D	1251 SE	2729 NI	1302 HE	2789 KT	1355 IT	2843 AL	1395 EP	2902 NS	1449 GS	2962 NI	1469 -	3022 00	1490 LE	3082 F	1527 PV	3142 VA	1573 SI	3197 AF	3249 KE	1677 AA	
ପ୍ର	∂ 1	<u> </u>	Š €	3 è	Š á	<u> </u>	Š	₹ .	음 원	ò	ģ	ò	q _Q	ò	셤	ò	셤	ò	ପ୍ଧ	ò	ପ୍ର	č	අධ	ò	q	ò	a	ò	d d	<i>&</i> ∙	do do	<u>ද</u>	ò	·
Query Match 3.0%; Score 399; DB 1; Length 4903; Best Local Similarity 17.9%; Pred. No. 1.1e-05; Matches 545; Conservative 345; Mismatches 1079; Indels 1080; Gaps 128;	90 SYLPELGKSEMEFIESKRPRLELLPDPLLRPSPLLATGQPAGSEDLTKDRSLTGKL 145	1458 SSLPQPGVSQSSRPLTEEQLDGILSPELDKMVTDGALLGKLYKIP 1502	146EPVSPPHTDPELELVPPKLSKEELIQNMDRVDREITM 185	1503 ELGGKDVEDLFTAVLSPATTQPAPLPQPPPPPQLLPMHNQDVFSRMPLMNGLI 1555	186 VEQQISKLKKKQQQLEEEAAKPPEPEKPVSPPFIESKHRSLVQIIYDENRKKAEAAHRIL 245	1556GPSPHLPHNSLPPGSGLGTFPAIAQSPYTDVRDKSPAFNAIA 1597	246 BGLGPQVBLPLVNQPSDTRQYHBNIKINQAMRKKLILLYFKRRNHARKQWKQKPCQ 300	1598 SDPNSSWAPTTPSMEGENDTLSNAQRSTLKWEKEEALGEMAT 1639	301RYDQLMBALEKKVERIENNPRRAKESKVREYYEKQFPEIRKQRELQERMQ 351	1640 VAPVLYTNINFPNLKEEFPDWTTRVKQIAKLWRKASSQERAPYVQKARD 1688	SRVGQRGSGLSMSAARSEHEVSEIIDGLSBQENLEKQMRQLAVIPPMLYDADQQRIKFIN	NKAALKINKVQMSNBSMKRQQQQDSIDPSSRIDSD	*12 migliandphina andronnimms roland-frankrightan follandring 10	REPRESENT MAINTENAMENT MAINTENA	:	KEBEKPEVENDKBDLLKEKTDDTSGEDNDEKEAVASKGRKTANSOGRRKGRI		TRSMANEANSERAITPOOSARIASMEINESSRWITEREMETAKKGIJEHGRNWSAIAP	OESLSOSONSOPP-SPOMFSPGSSHSRPPSPVDPYAK	MVGSKTVSOCKNPVPNVKKPONI.DRII.OOHKI.KMEKEBNABBKKKKADAaas	MYTTREPREDECTION OF THE CONTRACT OF THE CONTRA	REAAFD	DPXAKPPDTPRPMTDOFSKPPSLPRSPVISEOSTKGPLTTGTSDHFTKPSPRTDAFORO	WITH BALHAS GINE V PROBECSGDATAN	:: :: ::	NSSDTESIPSPHTEAAKDTGONGPKPPATLGADGPPPGPPTP		PRRTSRAPIEPTPASEATGAPTPPPAP-PSP	SYSQTSGTARSNPDPYSQPPGTPRPNTIDPYSQQPPTPRPSPQTDMFVSSVANQRHTDPY	816		835PPVBEGEBQKPPAABELAVDTGKAEBPVKSECTEBABEGPAKGKDABAABATABGA 890	2257 PLIRPPDICSQIPRPPGPGRIDIFTHASSSAVRDPYDQPPVTPRPHSESFGTSQVV 2312	891LKAEKKEGGSGRATTAKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPT 948
Q M E	ò	QQ	ò	qa	ò	qa	ò	q	ð	q	ð i	a d	S 8	Š	3 A	ઠે	3 8	l è	3 8	.	; f	ે કે	3 8	Š	; <u>a</u>	ò	: <u>a</u>	Š	; A	ò	a	ò	qq	ò

qa	
ò	949 GDPRANASPQKPLDLKQLKQRAAAIPPIQVTKVH- 982
qq	2357 GTDTQNTVNMSQADTEKLRQRQKLREIILQQQQQKKIASRQEKGPQDTAVVPHPVPLPHW 2416
ò	983 BPREDAAPTKPAPPAPPPPQNLOPESDAPQ 1013
qq	2417 OPESINQAFTRPPPYPGSTRSPVIPPLGPRYAVFPKDQRGPYPPEVAGMGMRPHGFRFG 2476
ò	QPGSS
q	2477 FPGAGHGPMPSQDRFHVPQQIQGSGIPPHIRRPMSMEMPRPSNNPPLNNPVGLPQHFPPQ 2536
ờ	1061 BVLPLGLHDTARP 1094 :
q	2537 GLPVQQHNILGQAFİELRHRAPDGRSRLPFAASPSSVIESPSHPRHGNFLPRP 2589
ò	1095 VLPRP
Q	2590 DFPGPRHTDPIRQPSQCLSNQLPVHPNLEQVPPSQQEQGHPAHQSSIVMRPLNHPLS 2646
δ	AKAPVGPVTMGLPLPMDPKKLAPFSGVK
d .	2647 GEFSEAPLSTSTPAETSPDNLEIAGQSSAGLEEKLDSDDPSVKEL 2691
δ	1193 SVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYKGTITRIIGED 1250
q	2692 DYKDLEGVEVKDLDDEDLENIANIDTEDGKGDDLDTLD 2728
ò	1251 SPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPP 1301
අු	2729 NETNDPNLDDLLRSGEFDIJAYTDPELDLGDKKSMFNEELDLNVPIDDKLDNQCASVEP 2788
ζ	1302 HETAAPKRTYDMMEGRVGRAISSASIEGLMGRAIPPERHSPHHLKEOHHIRGS 1354
đ	2789 KTRDQGDKTWVLEDKDLPQRKSSVSSEIKTEALSPYSKEEIQSEIKNHDDSRGD 2842
δ	1355 ITQGIPRSYVBAQEDYLRREAKLLKREGTPPPPPPSRDLT 1394
ą	2843 ADTACSQA-ASAQTNHSDRGKTALLTTDQDMLEKRCNQENAGEVVSAIQGSTPLPARDVM 2901
ò	1395 BAYKTQALGPLKLKPAHEGLVATVKEAGRSIHEIPREELRHTPELPLAPRPLKE 1448
q	2902 NSCDITGSTPVLSSLLSNEKCDDSDIRPSGSSPPSLPISPSTHGSSLPPTLIVPPSPLLD 2961
ò	1449 GSITQGTPLKYD1468
qq	2962 NTVNSNVTVVPRINHAFSQGVPVNPGFIQGQSSVNHNLGTGKPTNQTVPLTNQSSTMSGP 3021
ò	1469KKHDVRSLIGSPGRIFPPVHP 1489
qq	3022 QQLMIPQTLAQQNRERPLLLEEQPLLLQDLLDQERQEQQQQRQMQAMİRQRSEPFFPNID 3081
ò	35GG
q	3082 FDAITDPIMKAKMVALKGINKVMAQNSLGMPPMVMSRFPFMGPSVAGTQNNDGQTLVPQA 3141
È	1527 PVIVPELGKPRQSPLTYEDHGAPPAGHLPRGSPVIMREPIPRL
đ	3142 VAQDGSITHQISRPNPPNFGPGFVNDSQRKQYEEWLQETQQLLQMQQKYLEEQIG 3196
<i></i> ∂.	1573 SLSSSKASQDRKLTSTPREIAKSPHSTVPEHHPHPISPYEHLLRGVSGVDLY 1624
qa j	3197 AHRKSKKALSAKQRTAKKAGREFPEEDAEQLKHVTEQQSMVQKQLEQIRKQQ 3248
ò	3 1625 RSHIPLARDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYP-PYLIRGYPDT 1676
qq	3249 KEHAELIEDPGA 3290
ò	1677 AALENRQTIINDYITSQQMHHNTATAMAQRADMLRGLSPRESSLALNYAAGPRG 1730

ď

"Cloning and characterization of an ATBF1 isoform that expresses in neuronal differentiation-dependent manner."; J. Biol. Chem. 270:26840-26848(1995).

SEQUENCE FROM N.A. (ISOFORM B).

TISSUE-Lung;
WEDLINE=96070776; PubMed=7592926;
Miura Y., Tam T., Ido A., Morinaga T., Miki T., Hashimoto
Miura Y., Tam T., Ido A., Morinaga T., Miki T., Hashimoto
Tamaoki T.;

SEQUENCE FROM N.A. (ISOFORM A)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;

```
3869
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3923 AGVLVSHEVARALGPKPPQLPFRPQDDLLARALAQGPKTVDVPASLPTPPHNNHEBLRIQ 3982
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3983 DHYGDRDTPDSFVPSS------SPESVVGVEV----NKYPDLSLVKEEPPE 4023
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3970 AMYSSSDSFTHLKQQNNLSNPPTPPASLPP-----TPPPWACQKWANGFATTEELAGK 3922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3784
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ...---kPEGTLSSDESSTKDGKLIEKQSPAEGLQT 3814
                                                                                                                                                                                                                                                                                                                                      ----pappasasdphrektqskpps--iqelelrslgyhgssyspegvepvspsss 2064
                                                                                                                                                                                                                                                                                                                                                                                                        3657 EPVPVPTPNVSAGQPCIESENKLPNSEFİKETSNQOTHVNAEADKPS------VETPNK 3709
                                                                                                                                                                                                                                                                                                                                                                                                                                    THDKGLPKHLEBLDKSHLEGELRPKQPGPVK--LGGEAAHLPHLRPLPESQPSSSPL-LQ 2121
                                                                                                                                                                                                                                                           -----AATFPPA-THCPL-----GGTLDGVYPTLME--PVLLPKEAPRVARPERPR 1952
                                                                                 PLSPGGPTHLTKPTTTSSSERER-----DRDRERDRDREREK-----SILTSTTTVEH 1823
                                                                                                           PFSESFOERERKERLREQOERORVOLMQEVDRORALQORMEMEQHCLMGAELANRTPVSQ 3445
                          -- PTPGTPATAMDRLAYLPTAPQPFSSRHSSS 1775
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2396 SPGGGGKAKVSGRPSSRKAKSPAPGLASGDRPPSVSSVHSEGDCNRRTPLTNRVWEDRPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2289 EINKKLNTHNRNEPEYNISQPGTEIFNMPAITGTGLMTYRSQAVQEHASTNMGLEAI---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ----IRKALMGKYDQWEESP--PLSANAFNPLNASASLPAAMPITAADGRSDHTLT
                                                                                                                                                                                                                                                                                                                                                                                                                                                           TAPGVKGHQRVVTLAQHI----SEVITQDYTRHHPQQLSAPLPAPLYSFPGASCPVLD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | | | :: | :|:
TYKGDTGNE----LLKHLLKNKKASSLLTQ---------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2176 LRRPPSDLYLPPDDHGAPARGSPHSEGGKRSPEPNKTSVLGGGEDG--IEPVSPPEGMTE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2234 PGHSRSAVY-----PLLYRDGEQTEPSRMGSKSPGNTSQPPAFFSKLTESNSAMVKSKKQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----AHLPLNPPRIQPPIAQLSLKTCTPAPGTVSSANPQ-----NGPPPRVEFDDNN
                                                                                                                                                                      3446 MPFY------GSDRPCDFLQPPRPLQQSPQHQQQIGPVLQQQNVQQG
                                                                                                                                                                                                   1864 ---SPRTQDALQ---QR----PSVLHN----TGMKGIITAVEPSKPTVLRSTSTSSPVRP
                                                                                                                                                                                                                                                                                                                    ADTGHAFLAKPPARSGLEPASSPSKGSEPRPLVPPVSGHATIARTPAKNLAPHHASPDP-
                                                                                                                                          1824 APIWRPGTEQSSGSSGGSGSGSRPA-----SHSHAHQHSPI-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SAGSTPFPYNPLI-------MRLQAGVMASPPPPGLPAGSGPLAG 2493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4024 PVFSPIIFILPSISGKNSESRRNDIKTEPGTLFFTSPFG-SSPNGPRSG 4071
TSLIMSQP--NFPMVPQQLQHQQHTAVISGHTSPARMPSLPGWQSNSAS-
                                1731 IIDLSQVPHLPVLVP-
                                                                                                                                                                                                                                                                                                                                                          3607 KD-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3710
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2346
                                                                                                                                                                                                                                                                                                                                                                                                                                               2065
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3785
                                                                                                                                                                                                                                       3487
                                                                                                                                                                                                                                                                                                                                                                                      2012
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3759
                                                                                                                                                                                                                                                                    1910
                                                                                                                                                                                                                                                                                                                              1953
                                                                                          1776
                                                                                                                        3386
                                                               3338
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                              g
                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                   쉱
                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                       ð
                                                          g
                                                                                          8
                                                                                                                                                   ò
                                ò
```

SEQUENCE OF 1-1190 FROW N.A. (ISOFORM A).
MEDLINE-99425270; PubMed=10493829;
Loftus B.J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R.,
Loftus B.J., Mason T., Crosby M.L., Barnstead M., Cronin L.,
Deslattes Mays A., Cao Y. Xu R.X., Kang H.-L., Mitchell S.,
Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;
Genome duplications and other features in 12 Mb of DNA sequence from human chromosome 16p and 16q.";
Genomics 60:295-308(1999).

TISSUE=Hepatoma;
MEDLINE=92049333; PubMed=1719379;
MOTINE=92049333; PubMed=1719379;
MOTINGAT T., Yasuda H., Higashio K., Tamaoki T.;
Motinaga T., Yasuda H., Higashio K., Tamaoki T.;
Motinaga T., Yasuda H., Higashio K., Tamaoki T.;
Motinaga T., Yasuda H., Higashio K., Tamaoki T.;
Mol. Cell. Biol. 11:6041-6049(1991).

core

ROJENCE OF 1151-3703 FROM N.A.

Kozlowicz A., McQuerry Y., Hotic M.;

Submitted (UUN-1998) to the EMBL/GenBank/DDBJ databases.

-!- FUNCTION: Transcriptional activator that binds to the AT-rich sequence of the enhancer element of the AFP gene.

-!- SUBUNT: Interacts with FNBP3 (By similarity).

-!- SUBCELLULAR LOCATION: Nuclear.

-!- ALTERNATIVE PRODUCTS:

Event=Alternative splicing; Named isoforms=2; Name=A;

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GO; GO:0005634; C:nucleus; TAS.
GO; GO:0003705; F:RNA polymerase II transcription factor acti. . .; TAS.
GO; GO:0006355; P:regulation of transcription, DNA-dependent; TAS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IsoId=015911-2; Sequence=VSP_006825; SIMILARITY: Contains 4 homeobox domains
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IsoId=Q15911-1; Sequence=Displayed;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           entities requires a license agreement (Se
or send an email to license@isb-sib.ch)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; L32832; AAC14462.1; ---
EMBL; D10250; BAA01095.1; ---
EMBL; AC002044; AAC31674.1; ---
EMBL; AC004943; AAC79153.1; ---
HSSP; P20263; 10CP.
TRANSPAC; T00168; ---
TRANSPAC; T01665; ---
Genew; HGNC;777; ATBFI.
MIM; 104155; ---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR001356; Homeobox.
InterPro; IPR001087; Znf_C2H2.
Pfam; PF000046; homeobox; 4.
Pfam; PF000096; zf-C2H2, 18.
ProDom; PD000010; Homeobox; 4.
PROSITE; PS00027; HOMEOBOX, 1; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Name=B;
```

ABEL HUMAN STANDARD; PRT; 3703 AA.

(015911; 015101; 013719;
16-0CT-2001 (Rel. 40, Created)
16-0CT-2001 (Rel. 40, Last sequence update)
10-0CT-2003 (Rel. 42, Last amnotation update)
Alpha-fetoprotein enhancer binding protein (AT motif-binding factor)
AT-binding transcription factor 1).

ABF1_HUMAN ID ABF1 H AC Q15911 DT 16-0CT DT 16-0CT DT 10-0CT DE Alpha-DB (AT-bh GN ATBF1.

RESULT

```
667 NPSS-----AGIAPSSNLLSQVPTESVGMPP-----LGNPIGA-- 1699
                                                                                                                                                                                                                                            1449 ETSHLELSEADIQQLYGGLLANGDLLAMGDPTLAEDHTIIVEEDKEEESDLEDKQSPTGS 1508
                                                                                                                                                                                                                                                                                                                                                  1509 DSGSVQEDSGSEPKRALPFRKGPNFTMEKFLDPSRPYKCTVCKESFTQ-----KNILLV 1562
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      --SLSSSTPSPVSTSGSNTFTTS 1666
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | ||: :| ||
2150 TRITDDQLRVLRQYFDINNSPSEEQIKEMAD----KSGLPQKVIKHWFRNTLFKERQRNK 2205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2206 DSPYNFSNPPITSLBELKIDSRPPSPEPPKQEYWGSKRSSRTRFTDYQLRVLQDFFDANA 2265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1895 SAEGG---EG-----AKE 1917
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2090 MPMELPIFSPLMMQTMPLQTLPAQLPPQLGPVEPLPADLAQLYQHQLNPTLLQQQNKRPR 2149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2266 YPKDDĖFEQLSNLINLPTRVIVVWFQNARQKARKNYĖNQGEGKDGERRĖLINDRYIRTSN 2325
                                                                                                                                                                                                                                                                                                --PPHIDPELELVPPR----LSKEELIQNMDRVDREITMV 186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           366 ARSEHEVSEIIDGLSEQENLEKQMRQLAV-IPPMLYDADQQRIKFINMNGLMADPMKVYK 424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   KKKKAPAAASEEAAFPPVVEDEEMEASGVSGN----------EEEMVEEAE 722
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -PEVENDKEDLLKEKTD------DTSGEDNDEKEAVASKGRKTANSQGRRKGRITR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1972 GKTDQGENLEKLECDSCGKLFSNILILKSHQEHVHQNYFPFKQLERFAKQYRDHYDKLYP
                                                                                                                                                                                                                                                                                                                                                                                                    187 E-QQISKLKKKQQQLEEEAAKPPEPEKPVSPPPIESKHRSLVQIIYDENRKKAEAAHRIL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  306 MEALEKKVERIENNPRRRAKESKVREYYEKQFPEIRKQRELQERMOSRVGQRGSGLSMSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              425 DROVMNAMSEQE-KETFREKFMQHPKNFGLIASFLERKTVAECVLYYYLTKKNENYKSLV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      484 RRSYRRRGKSQQQQQQQQQQQQQQQQPMPRSSQBEKD--EKEKEKEAEKEEEK-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1777 LPHFPMTTETLLQLQQQQHLLFPFYIPSAEFQLNPEVSLPVTSGALTLTGTG--PGLLED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TAKKGLLEHGRNWSA I ARMVGSKTVSQCKNFY FNYKKRQNLDE I LQQHKLKMEKERNARR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            723 ALHASGNEVPRGECSGPATVNNS----SDTESIPS---PHTEAKDTGQNG-----P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 768 KPPATLGADGPPPPPPPPRRTSRAPIEP--TPASEATGAP--TPPPAPPSPSAP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -- EGPAKGKDAEAAEATAEGALKAEK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EGLGPQVELPLYNQPSDTRQYHENIKINQAMRKKLILYFKRRNHARKQWKQKFCQRYDQL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----PKEEKEETAAAPPVEEGEEQK-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----PP--AABELAVDT-GKABEPVKSEC-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ----NIASPSEPKEANRKKLAD-----MIAS-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----YSQ-SSTLEIHMRSVLHQT-----YSQ
                                                                                                                                                                                                                                                                                                                                                                                                                                     1623 -KARAAKLEAASGSSNGTGNSSSI------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ---TEEAE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1605
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             622
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1700
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       819
                                                                                                                                                                                                                                                                                                153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1722
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               536
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   682
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ..... 867
                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                 ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ક
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          원
HOMEOBOX 2; 4.

.J28; ZINC_FINGER_C2H2_1; 14.
.S50157; ZINC_FINGER_C2H2_2; 9.
.iption regulation; Activator; Zinc_finger; Metal-binding;
.Lernatus splinding. Activator; Zinc_finger; Metal-binding;
.Lernatus splinding. Activator; Zinc_finger; Metal-binding;
.Lernatus splinding. Activator; Zinc_finger; Metal-binding;
.ZN FING 282 305 C2H2_TYPE.
FT ZN FING 671 694 C2H2_TYPE.
FT ZN FING 671 694 C2H2_TYPE.
FT ZN FING 945 968 C2H2_TYPE (ATYPICATOR)
FT ZN FING 1040 1064 C2H2_TYPE (DEGTTAL)
FT ZN FING 1040 1064 C2H2_TYPE (FT ZN FING 122) 1246
FT ZN FING 122 1246
FT ZN FING 1360 1385
FT ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FING 1401
T ZN FIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AIEDVEGPSETAADPEELAKDQEGGASSSQAEKELTDSP
GEWSHRHGRPRLGLGVHLLETSRGLLPEGDVTDPAGPH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 3.0%; Score 396; DB 1; Length 3703; Similarity 18.4%; Pred. No. 1.1e-05;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /FTId VAR 011696.
A -> S (IN REF. 3).
A -> P (IN REF. 3).
T -> A (IN REF. 3).
T -> A (IN REF. 3).
A -> V (IN REF. 3).
HHRV -> RHLG (IN REF. 3).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3703 AA; 404468 MW; 0F62AF37D4DCF856 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Missing (in isoform B)
/FTId=VSP_006825.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        C2H2-TYPE (ATYPICAL).
C2H2-TYPE.
HOMEOBOX 3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A -> V.
/FTId=VAR_011694.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /FTId=VAR_011695.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A -> S (IN REF.
EEAIEDVEGPSETAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HOMEOBOX 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -> GEWSHRHG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HOMEOBOX 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            C2H2-TYPE.
C2H2-TYPE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        C2H2-TYPE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9999 ^-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POLY-SER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Missing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Missing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
POLY-I
PO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                72
422
579
767
777
849
997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2301
2351
22552
2736
2736
2736
3553
3553
3553
489
784
784
11794
11794
11794
11794
32052
32052
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3389
33395
3527
3600
3639
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3374
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3507
3597
3636
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3374
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     037
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ZN_FING
DNA_BIND
DNA_BIND
ZN_FING
ZN_FING
DNA_BIND
ZN_FING
DNA_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ZN FING
DNA BIND
ZN FING
ZN FING
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /ARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
```

KEGGSGRATTAKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRANA 955

ò

Gaps 115;

EMEFIESKRPRLELLPDPLLRPSPLLATGQPAGSEDLT----KDRSLTGKLEPVSPPS-- 152

= = = =

Conservative 271; Mismatches 968; Indels 1036;

Matches 514;

		à	1905 SPVRPA
ờ	956 SPQKPLDL-KQLKQRAAAIPPIQVIKVHEPPREDAAPTKPAPPAPPPP 1002	· £	3214PAZ
g	2386 MPSQAYSAPAPSANNTASSAFLQLTAEAEBLATFNSKTEAGDEKPKLAEAPSAQPNQTQE 2445	ł	40
ò	1003 ONLOPESDAPOOPGSSPRGKSRSPAPPADKEAFAABAQKLPGDPPCWTSGLPFPVPPREV 1062	ें ह	
gg	: : : : : : : : : : : : :	3 8	
è	1063 IKASPHAPDPSAFSYAPPGHPLPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLER 1120)	
ÿ ;		g	3312 VPGFSP
Q	AND THE STATE OF T	ò	2084 GELRPK
ờ	QIGAISQGMSVQLHVYYSE-HAKAFVGFVIMGLFLFFMLFANDFT	ą G	3356
අු		δ	2144 TQDYTR
λo		đ	3389 -QKVQQ
<u>ප</u> ු .		ò	2204 KRSPE-
ò	KKGHVL ::	QQ	3425 KNTPRE
අු	STWNTLKRKLEEKASASPGENDSGTGGEEPQRDKRLRTTITPEQLEILYQKYLL	ò	2242 YPLLYR
ò	ISSASIEGLMG	QD	3485 KSLCFF
ą	2663 DSNPTRKMLDHIAHEVGLKKRVVQVWFQNTRARERKGQF 2701	ò	2287 KQEINK
ζ		QQ	3545 SQHLES
qa	2702 RAVGPAQAHRRCPFCRALFKAKTALEAHIRSRHWHEAKRAGYNLTLSAMLLDCDGGLQMK 2761	ò	2330 -QAVQE
δ	1353 GSITQGIPRSYVEAQEDYLRREAKLLKREGTPPPPPPSRDLTEAYKTQALGPLKL 1407	. d	
qa	2762 GDIFDGTSFSHLPPSSSDGQGVPLSPVSKTMELSPRTL 2799	à	2380 -AMPIT
δλ	1408 KPAHEGLVATVKEAGRSIHEIPREELRHTPELPLAPRPLKEGSITQGTPLKXDTGASTTG 1467	q	: : 3653 PSMPTD
qq	2800 LSPSSIKVEGIEDFESPSMSSVNLNFDQTKLDNDCSSVNTAITDTTGDEG 2851		
ζ	1468 SKKHDVRSLIGSPGRTFPPVHPLDVMADARALERACYEESLKSRPGTASSSGGSIARGAP 1527	RESULT TCOF H	RESULT 32 TCOF_HUMAN
QQ	2852 NADNDSATGIATETKSSSAPNEGLIKAAMMAMSEYEDRLSSGLVSPA 2898	AC D	TCOF HUMAN Q13428; Q99408;
ò	1528 VIVPELGKPROSPLTYEDHGAPFAGHLPRGSPVTWREPTPRLQEGSLSSSKASODR 1583	TO	01-NOV-1997 (Re 01-NOV-1997 (Re
qq	2899PSFYSKEYDNEGTVDYSETSSLADPCSPSPG-ASGSAGKSGDSGDRPGQK 2947	D D	u
ò	1584KLTSTPREIAKSPHSTVPEHHPHPISPYEHLLRGVSGVDLYRSHIPLAFDPTSIPR 1639	GN	TCOF1.
Q	2948 RPRIQMINIQLKVLKSCFNDXRIPIWLECEVLGNDI 2983	888	Eukaryota; Meta
ò	1640 GIPLDAAAAYYLPRHLAPNPT-YPHLYPPYLIRGYPDTAALENRQTI 1685	X X X	NCBI_TaxID=9606
q	: : : :	RR	SEQUENCE FROM N MEDLINE=9615418
ò	1686INDYITSQQMHHNTATAMAQR-ADMLRGLSPRESSLALNYAAGPR 1729	RA 8	Dixon J., Edwar
q	3044 FSQQHISKVKDTIGSQLDKEKEYFDPATVRQLMAQQELDRIKKANEVLGLAAQQQ 3098	RT	"Positional clo
ò	1730 GIIDLSQVPHLPVLVPPTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTH 1784	R. R.L.	Nat. Genet. 12:
qq	3099 GMFDNTPLQALNLPTAXPALQGIPPVLLPGLNSPSLPGFTPSNTA 3143	RP	SEQUENCE FROM NEDLINE=9722890
ò	1785 LIKPTITSSSERERDRDRERERDRDREREKSILISTITVEHAPIWRPGIEQSSGSSGSGGG 1844	8 8	Dixon J., Edwardixon M.J.;
q	3144 LISPKPNIMGLPSTTVPSPGL-PTSGLPNKPSSA 3176	RT	"Identification organization of
ò	1845 GGSSSRPASHSHAHQHSPISPRIQDALQQRPSVLHNTGMKGIITAVEPSKPTVLRSTSTS 1904	RL RN	Genome Res. 7:2
qu	3177 SLSSPTPAQATWAMGPQQPPQQQQQQQQQQQQQQVQQPPPP 3213	RX	SEQUENCE FROM NEDLINE=9725049

EHASTNMGLEAIIRKALMGKYDQWEESPPLSANAFNPLNASASLPA----- 2379 AATFPPATHCPLGGTLDGVYPTLMEPVLLPKEAPRVARPERPRADTGHAFLAKPP 1964 SPASSPSKGSEPRPLVPPVSGHATIARTPAKNLAPHHA-SPDPPAPPASASDPHR 2023 KPFSIQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDKGLPKHLEELDKSHLE 2083 KOPGPVKLGGEAAHLPHLRPLPESQPSSSPLLQTAPGVKGHORVVTLAQHISEVI 2143 -SPGSL-----LQQYQQYQQSLQEAIQ-----QQQQQQQQQ----- 3388 AHHPQQLSAPLPAPLYSFPGASCPVLDLRRPPSDLYLPPPDHGAPARGSPHSFGG 2203 ------PNKTSVLGGGEDGI-EPVSPPE------GMTEPGHSRSAV 2241 RDGEQTEPSRM-----GSKSPGNTSQPPAFFSKLTESNSAMVKSK 2286 FGQSVVNLQEMVLHVPTGGGGGGGGGGGGGGGGGGGGSYHCLACESALCGEEAL 3544 KKLNTHN-----RNEPEYNISQPGTEIFNMPAITGTGLMTYRS----- 2329 PYYAPQIPGALQSGYLQPWYGMEGLFPYSPALSQAL---MGL------.azoa. Chordata, Craniata, Vertebrata, Buteleostomi, neria, Primates, Catarrhini, Hominidae, Homo. 183; PubMed=8563749; ards S.J., Gladwin A.J., Dixon M.J., Loftus S.K., Koprivnikar K., Wasmuth J.J.; loning of a gene involved in the pathogenesis of ins syndrome."; 2:130-136(1996). N.A. 300; PubMed=9074926; ards S.J., Anderson I., Brass A., Scambler P.J., on of the complete coding sequence and genomic of the Treacher Collins syndrome gene."; 233-234(1997). Rel. 35, Created)
Rel. 35, Last sequence update)
Rel. 41, Last annotation update)
in (Treacher Collins syndrome protein). PRT; 1411 AA. TAADGRSDHTLTSPGGGGKAKVSG 2407 DDYSEESDIDLSOKSDGPASPVEG 3681 N.A. 498; PubMed=9096354; STANDARD; ; Q99860; Human)

```
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                   Am. J. Hum. Genet. 60:515-524(1997).

-!- DISEASE: Defects in TCOP1 are the cause of Treacher Collins syndrome (TCS) [MIM:154500]. TCS is an autosomal dominant disorder of craniofacial development that occurs with an incidence of 1/50,000 live births. The clinical features of TCS are bilaterally symmetrical and include: (1) abnormalities of the external ears, atreata of the external ear canals, and malformation of the middle ear ossicles, which may result in conductive hearing loss; (2) lateral downward sloping of palpebral fissures, frequently with colbomas of the lower eyellds; (3) hypoplasia of the mandible and zygomatic complex; (4) cleft palate.

-!- SIMILARITY: Contains 1 Lish domain.
                                                      exhibits
                                                                                                                                                                         VARIANTS LEU-439; VAL-810; VAL-1313 AND GLY-1355, AND VARIANT TCS
                                                                                                                                                                                                                                       Edwards S.J., Gladwin A.J., Dixon M.J., The mutational spectrum in Treacher Collins syndrome reveals predominance of mutations that create a premature-termination
Wise C.A., Chiang L.C., Paznekas W.A., Sharma M., Musy M.M., Ashley J.A., Lovett M., Jabs E.W.; "TCOF1 gene encodes a putative nucleolar phosphoprotein that mutations in Treacher Collins syndrome throughout its coding
                                                                                                                       Proc. Natl. Acad. Sci. U.S.A. 94:3110-3115(1997)
                                                                                                                                                                                          ARG-53.
MEDLINE=97195537; PubMed=9042910;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             JOINED.
JOINED.
JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              JOINED.
JOINED.
JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL, U40847, AAC50903.1, EMBL, U76366, AAC51181.1; EMBL, U84664, AAC51185.1, COINE EMBL, U84641, AAC51185.1, JOINE EMBL, U84641, AAC51185.1, JOINE EMBL, U84642, AAC51185.1, JOINE EMBL, U84644, AAC51185.1, JOINE EMBL, U84646, AAC51185.1, JOINE EMBL, U84646, AAC51185.1, JOINE EMBL, U84649, AAC51185.1, JOINE EMBL, U84649, AAC51185.1, JOINE EMBL, U84649, AAC51185.1, JOINE EMBL, U84649, AAC51185.1, JOINE EMBL, U84649, AAC51185.1, JOINE EMBL, U84649, AAC51185.1, JOINE EMBL, U84649, AAC51185.1, JOINE EMBL, U84650, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U84651, AAC51185.1, JOINE EMBL, U846
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAB40722.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAB40722.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAB40722.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAB40722.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAC51185.1;
AAC51185.1;
AAC51185.1;
AAC51185.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AACS1185.1
AACS1185.1
AACS1185.1
AACS1185.1
AACS1185.1
AACS1185.1
AACS1185.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAB40722.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAB40722.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAB40722.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAB40722.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAB40722.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               U79654;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        U84653;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   U84654;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       384655;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      U84662;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                J84656;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             J84658;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            384660;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    384663;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         U79648;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      U84657
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     JB4661
                                                                                                                                                                                                                                                                                                                  codon."
                                                                                                      region.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
EMBL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL;
```

```
958 QKPLDLKQLKQRAAAIPPIQVTKVHEPPREDAAPTKPAPPAPPPPQNLQPESDAPQQPGS 1017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1018 SPRGKSRSPAPPADKEAFAA--EAQ----KLPGDPPCWTSGL-PF----PVPPREVIK 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   707
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                50 YTHWQQTSELGRKRKABEDAALQAKKTRVSDPISTSESSE----EEEEAEAETAKATPR 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ----SGVSGNE-----EEMVEEAEALHASGNEVPRGECSGPAT-----VNNSSDTE 749
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               105 LASTNSSVLGADLPSSMKEKAKAETEKAGKTGNSMPH-----PÄTGKTVANLLSGKSPRK 159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----RISRAPIEPTPASEATGAPTPPPAPPSPSAPPPVVPKEEKEETAAAPPVEEG 840
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         217 SEKILQVRAASAPAKGTPGKGAT-----PAPPGKAGAVASQTKAGKPEEDSESSSEESS 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          E-EQKPPAAEEL--AVDTGKAEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKE 897
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     271 DSEBETPAAKALLOAKASGKTSO--VGAASAPAKESPRKG------AAPAPP 314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GGSGRATTAKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRANASP 957
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ----- PAQAKPSG----- 354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      355 -----PRRYTGPAAGAP-----AKESPRKGAAPA------PRKYTGPAAQVQVGKQ 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                395 EEDSRSSSEESDSDREALAAMNAAQVKPLGKSPQVKPASTWGWGPLGKGAGPVPPGKVGP 454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | || :| :| :| | | | | | | SAEPSANTTLVSETEEEGSVP--AFGA-AAKPGMVSAGQADSSSEDTSSSSDETDVEVKA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   656 YKKRONLDEILQQHKLKMEKERNARRKKKKAPAAASEEAAFPPVVEDEEMEA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 3.0%; Score 391.5; DB 1; Length 1411; Best Local Similarity 20.4%; Pred. No. 6.1e-06; Matches 364; Conservative 178; Mismatches 664; Indels 581; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 1411;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               750 SI-PSPHTEAAKDTGQNGPKPPATLGADGPPPGPPTPPR------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /FTId=VAR 005634.
K -> Q (IN REF. 2).
MW; 3880203D985C2699 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A-> v. -
/FTId=VAR_005632.
A-> v (in dbSNP:15251).
/FTId=VAR_005633.
                                                                                                                                                                                                                                                                                                                                                                                                     LISH.
POLY-GLU.
POLY-SER.
POLY-SER.
POLY-SER.
POLY-LYS.
POLY-LYS.
POLY-LYS.
W -> R (in TCS).
/FTId=VAR_005630.
                                                                                                           MIM; 606847; -
MIM; 154500; -
GO; GO:0005710; C:nucleolus; TAS.
GO; GO:0005215; F:transporter activity; TAS.
GO; GO:0001501; P:skeletal development; TAS.
InterPro; IPR000559; LisH.
InterPro; IPR003993; treacle.
Pfam; PF03546; treacle, 3.
PRINTS; PR01503; TREACLE.
SMART; SW00667; LisH; 1.
PR051TE; PS50896; LiSH; 1.
PR051TE; PS50896; LiSH; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     P -> L.
/FTId=VAR_005631
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | : | : : | : : | : | : | 315 GKTGPAVAKAQAGKREEDSQSSSEESDSEEBA-
                      EMBL; U79656; AAB40722.1; JOINED.
EMBL; U79657; AAB40722.1; JOINED.
EMBL; U79658; AAB40722.1; JOINED.
Genew; HGNC:11654; TCOF1.
EMBL; U79655; AAB40722.1; JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1411 AA; 144312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1313
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          841
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  708
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       868
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
```

1234 PADVILYKGTITRIJGEDSPRILDRGREDSLPKGHVIYBGKKGHVLSYBGGONSVTOC 578 KVAPUWRUT 1

```
copyright. It is produced through a collaboration trute of Bioinformatics and the EMBL outstation. ics Institute. There are no restrictions on its stitutions as long as its content is in no way ment is not removed. Usage by and for commercial anse agreement (see http://www.isb-sib.ch/announce/ense@isb-sib.ch).
                                           RKGWESRKRKLSGDOPAA----RTPRSKKKKL------ 1290
                                                                                          2202
                                                                                                                                                                             HEDGIEPVSPPEGMTEPGHSRSAVYPLLYRDGEQTEPSR-----M 2257
                                                                                                                                                                                                        GGEAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQRVVTLAQHISEV 2142
                                                                                                                                   ------GA-----GEG 1295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MGC1114 10.34

MGC1114 10.34

MGC1114 10.34

MGC11A1 CDNA and identification of a

no.1agen Gly208Ala mutation in a severe case
s imperfecta.";
s. 384:37-46(200).
Goollagen is a member of group I collagen
collagen).
of one alpha 2(1) and two alpha 1(1) chains.
of one alpha 2(1) and two alpha 1(1) chains.
the third position of the tripeptide repeating
that catal a some or all of the chains.
in COLIA1 are a cause of osteogenesis imperfecta
                                                                                       1.PAPLYSFPGASCPVLDLRRPPSDLYLPPPDHGAPARGSPHSEG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 data, Craniata, Vertebrata, Buteleostomi, ivora, Fissipedia, Canidae, Canis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1; 1.
2; 1.
Jonnective tissue; Repeat; Hydroxylation;
Signal; Disease mutation.
SY SIMILARITY.
                                                                                                                                                                                                                                                                  KLTESNSAMVKSKKQEIN--KKLNTHNRNEP 2302
                                                                                                                                                                                                                                                                                         st sequence update)
st annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                           PRT; 1460 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             VARIANT OI ALA-208.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 VWFC domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Tagen.
collagen_C.
brinogen_C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ; 18.
elix; 2.
ollagen_C; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            n precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            =11147834;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     helix.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           eated)
```

1469 KRHD	SULT 34 K2 HUMAN ANKO HUMAN O1464, Q1485, 01.APR-1993 (Rel. 25, Created) 01.OCT-1996 (Rel. 34, Last sequence update 28.FEB-2003 (Rel. 41, Last sequence update Ankyrin 2 (Brain ankyrin) (Ankyrin B) (Ank	PLAKERPARSGLEPASSESRGERPLUVPPVSGHATIARTPAKALAPHHASPDPPAPASA GLOGPPGPPGSPGASGPAGPRGPPGSA GLOGPPGPPGSPGASGPAGPRGPPGBB GLOGPPGPPGSPGASGP
PROPER 135 1214 COLLAGAM ALPHA ICIT CHAIN 158 1214 COLLAGAM ALPHA ICIT CHAIN 135 1214 COLLAGAM ALPHA ICIT CHAIN 135 1214 COLLAGAM ALPHA ICIT CHAIN 136 1214 NUMPELICAL REGION (N-TERNINAL) 138 1214 NUMPELICAL REGION (C-TERNINAL) 138 1214 NUMPELICAL REGION (C-TERNINAL) 1315 1214 NUMPELICAL REGION (C-TERNINAL) 1315 1314 CELL ATTACHNET SITE (COTENITAL) 1315 1314 CELL ATTACHNET SITE (COTENITAL) 1315 1314 CELL ATTACHNET SITE (COTENITAL) 1315 1314 CELL ATTACHNET SITE (COTENITAL) 1315 1314 CELL ATTACHNET SITE (COTENITAL) 1315 1314 CELL ATTACHNET SITE (COTENITAL) 1315 1314 CELL ATTACHNET SITE (COTENITAL) 1315 1314 CELL ATTACHNET SITE (COTENITAL) 1315 1314 CELL ATTACHNET SITE (COTENITAL) 1315 1314 CELL ATTACHNET SITE (COTENITAL) 1315 COLDESTATION 1315 COLDESTATIO	Qy 1296 SSSGPPHETAAPKRTYDDMBGRVGRAISSASIEGLMGRAIPP-ERHSPHHIKEQHHIRGS 1354	975 PIQUTKVHEPPREDAAPTKPA-PPAPPPQNLQPESDAPQQPGS-SPRGKSRSPAPP 320 NDGATGAAGPPGPTGPAGPPGPPGAVGAKGBAGPQGARGBGPPGPA 1030ADKEAFAARAQKLPGDPPCWTSGLP-FPVPPREVIXASPHAPDPSAFSYA 376 GAAGPAGNPGAKGAMGAPGIAGAPGFPGARGPSGPGPGPRGNSG 1079 PPGHPLPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHVPYS 1

REVISIONS

```
DEATH.
              SMART; SM00248; ANK; 22.
SMART; SM00005; DEATH; 1.
SMART; SM00218; ZU5; 1.
PRINTS; PR01415; ANKYRIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    476
971
3582
3586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        822
1950
1784
1796
1808
1820
1832
1844
                                                                                                                          Phosphorylation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1785
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1845
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1809
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CONFLICT
                                                                                                                                                         REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                          REPEAT
                                                                                                                                                                       REPEAT
                                                                                                                                                                                       REPEAT
                                                                                                                                                                                                      REPEAT
                                                                                                                                                                                                                      REPEAT
                                                                                                                                                                                                                                     REPEAT
                                                                                                                                                                                                                                                                                                                                                                 REPEAT
                                                                                                                                                                                                                                                                                                                                                                               REPEAT
                                                                                                                                                                                                                                                                                                                                                                                            REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REPEAT
                                                                                                                                                                                                                                                                    REPEAT
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TISSUE SPECIFICITY: Distance membrane of neurons as well as glial cells throughout the brain.

PTM: Phosphorylated at multiple sites by different protein kinases and each phosphorylation event regulates the protein's structure and function (Potential).

SIMILARITY: Contains 1 death domain.
                                                                                                                                                                                                                                                                                                TISSUE=Brain stem;
MEDLINE=94075409; PubMed=8253844;
Chan W., Kordell E., Bennett V.;
"440-kD ankyrinB: structure of the major developmentally regulated domain and selective localization in unmyelinated axons.";
                             Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia, Eutheria, Primates, Catarrhini, Hominidae, Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE=92009921; PubMed=183308;
TSE W.T., Menninger J.C., Yang-Feng T.L., Francke U., Sahr K.E.,
Lux S.E., Ward D.C., Forget B.G.;
"Isolation and chromosomal localization of a novel nonerythroid
ankyrin gene.";
Genomics 10:858-866(1991)
-!- FUNCTION: Attach integral membrane proteins to cytoskeletal
elements. Also bind to cytoskeletal proteins.
                                                                                                      TISSUE=Brain stem;
MEDLINE=91302466; PubMed=1830053;
Otto E., Kunimoto M., McLaughlin T., Bennett V.;
"Isolation and characterization of cDNAs encoding human brain ankyrins reveal a family of alternatively spliced genes.";
J. Cell Biol. 114:241-253(1991).
                                                                                                                                                                                                                                   Carpenter S.; Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IsoId=Q01484-2; Sequence=VSP_000267, VSP_000268;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Event=Alternative splicing; Named isoforms=3,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Name=3;
IsoId=Q01484-3; Sequence=VSP_000268;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Name=1;
IsoId=Q01484-1; Sequence=Displayed;
                                                                                       SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
                                                                                                                                                                                                                                                                                                                                                                            Cell Biol. 123:1463-1473(1993).
                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A. (ISOFORM 1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; X56957; CAA40278.1; -.
EMBL; X56584; CAA40279.2; -.
EMBL; Z26634; CAA42644.1; -.
EMBL; M37123; AAA62828.1; -.
PIR, S37431, S37431. B287, B2771; 1DC2.
                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE OF 463-495 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR002110; ANK.
InterPro; IPR000488; Death.
InterPro; IPR000966; ZUS.
Pfam; PF00023; ank; 24.
Pfam; PF00531; death; 1.
Pfam; PF00791; ZUS; 1.
                Homo sapiens (Human).
                                                        NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MIM; 106410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Name=2;
```

```
Matches 543; Conservative 345; Mismatches 1023; Indels 1044; Gaps 129;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         942
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       140 SLTGKLEPVSPPSPPHTDPELELVPPRLSKEELIQNMDRVDREITMVEQQISKLKKKQQQ 199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        200 LEEEBAAKPPE--PEKPVSPPPIES------KHRSLVQIIYDENRKKAE 239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        83 HIRPESHSYLPELGKSEM---EFIESKRPRIELLPDPLLRPSPLLATGOPAGSEDLTKDR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      S------HT----HT-----LSHASYLRDSAVMDDSVVIPSHQVSTLAKEAER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     861 YLRPED---LKELGDDSLPSSQFLDGMN-----YLRYS--LEGGRSDSLRSFSSDR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     - > OFLGKLHLPTAPPPLNEGESLVSRILQLGPPGTK
in isoform 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /FIId=VSP 000267.
Missing (in isoform 2 and isoform 3).
/FIId=VSP 000268.
GQ -> PE (IN REF. 4).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3.0%; Score 391; DB 1; Length 3924; Similarity 18.4%; Pred. No. 1.6e-05;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3581 3582 QY -> HA (IN REF. 1).
3586 3586 I -> Y (IN REF. 1).
3924 AA; 430337 MW; 52AC496C428E29D2 CRC64;
PROSITE; PSS0088; ANK REPEAT; 20.
PROSITE; PSS0297; ANK REP REGION; 1.
PROSITE; PSS0017; DEATH DOMAIN; 1.
Cytoskeleton; Alternative splicing; Repeat; ANK repeat;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REPEAT A.
REPEAT A.
REPEAT A.
REPEAT A.
REPEAT A.
REPEAT A.
REPEAT A.
REPEAT A.
REPEAT A.
REPEAT A.
REPEAT A.
REPEAT A.
REPEAT A.
REPEAT A.
REPEAT A.
REPEAT A.
REPEAT A.
REPEAT A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              I -> S (IN REF. 1).
QY -> HA (IN REF. 1).
I -> Y (IN REF. 1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ANK 23.
REPEAT-RICH REGION.
REPEAT A.
```

01 Db 1901 EKRLPVSPSGRTDKG	Db 1960	<i>λ</i> ο ·	2009	Oy 1182 ESECVETAGEASVL O Dh 2023I.DVITGTAREKK	1242	Db 2044	1223 Qy 1302 HETAAPKRTYDMMEG		, Qy 1357	2139	ò		À i	2239	Qy 1518 SGGSIARGAPVIVPE	0 10 10 10 10 10 10 10 10 10 10 10 10 10	15/2 SSSKA	1628	Db 2336	91 Qy 1686 INDYITSQOMHHNTA	Db 2386	26 OY 1742 VLVPPTPGTPATAME	Db 2420 DSLEPSPLKESPCRE	Oy 1797 BRDRDR-ERDF	Db 2466 LTEVASVRSRLLRDF	Qy 1852	Db 2509 SEBVSYEUTPK		Db 2566 DELEQEAKOKR	Oy 7.1938 LPKEAPRVAR		Db 2674
:	240 AAHRILEG	NPRRRAKES		328 KVREYYEKQFPEIRKQRELQERMGSRVGQRGSGLSMSAARSEHEVSEIIDGLS 380	CRIITRDFPOYFAVVSRIKQDSNLIGPEGGVLSSTVVPQVQAVFPEGALTKRIRVGLQ		AQPMHSELVKKILGNKATFSPIV-TLEPRRRKFHKPITWTIPVPKASSDVMLNGFGGDAP		NFGLIA		456SFLERKTV-AECVLYYYLTKKNENYKSLVRRSY 487	: :: : : : : : : : :	488 RRRGKSGQQQQQQQQQQQQQQPMPRSSQEEKDEKEKEKEAEKEEEK-PEVENDKEDLL 546	: : : : : : : : : :	547 KEKTDDTSGEDNDEKEAVASKGRKTANSQGRRKGRITRSMANEANSEAIT 597	1455 KSHLVNBVPVLASPDLLSEVSEMKQDLIKMTAILTTDVSDKAGSIKVKELVKAAEEEPGE 15	598 PQQSABLASMELSSRWTEEBMETAKKGLLEHGRNWSAIARMVGSKTVS 647	1515 PREIVERVKEDLEKVNÄILRSGTCTRDESSVQSSRSERGLVEEBWVIVS 15	QCKNFYFNYKKRONLDEILLQQHKLKMEKERNARRKKKKADAAASBEAAFPPVVEDEE				SOURCE S	LAVDTGKARRDVKS	RRKLKEKOKOKEEGLOASAEKAELKKGSSEESLGEDPGIAPEPLPTVKATSPLIEETPIG		SIKDKVKALOKRVEDEQKGRSKLPIRVKGKEDVPKKTTHRPHPAASPSLKSERHAPGSPS	903 ATTAKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRANASPQK 959		960 PLDLKQLKQRAAAIPPIOVTKV-HEPPREDAAPTKPAPPAPPPPONLOPESDAPQQPGSS 1018	1847 PVSPSTKTERHSPVSSTKTERHPPVSPSGKTDKRPPVSPSGRTEKHPPVSPGRT 1900	1019 PRGKSRSPAPPADKEAFAAEAQKLPGDPCWTSGLPPPVPPRE 1061
qq	රි සි	ò	qq	ò	qa	È	ପ୍ର	පි සි	3 &	q	ò	q	ò	qq	ò	qq	ò	Ωρ	ð i	g	ර ස්	3 8	3 a	8	g 45	ð	; A	Š	qa	ŏ	qq	δ

PELGKPRQSPLTYEDHGAPFAGHLPRGSPVTMREPTPRLQE---GSL 1574 DS-----ETSTESFOREATLGSP---KDTSPKRQDDCTGSC 2298 :: BEAACDEGGRIFGSSAHKIQTDSE-----2335 | | | | : : | | : : | RDSL------ESSPVEPKWKAGIFPSHFPLPAAVAKTEL 2465 FSYAPPGHPLPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLERQ 1121 ----- AQKTENQTIKRĞQR--- 2022 LRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYKG 1241 LDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPP 1301 SGRVGRAISSASIEGLMGRAIPPERHSPHHLKE-----QHHIRGSIT 1356 -----AQEDYLRREAKLLKREGTPPPP-----PPSRDLTEAY 1397 HEGLVATVKEAGRSIHEIPREELRHTPELPLAPRPLKEGSITOGTPL 1457 HDVRSLIGSPGRTFPPVHPLDVMADARALERACYEESLKSRPGTASS 1517 -------SEKEHPTTKDI 2252 JORKLISTPREIAKSPHSTVPEHHPHPISPYEHLLRGVSGVDLYRSH 1627 IPLDAAAAYYLPRHLAPNPTYPHLYPPYLIRGYPD--TAALENRQTI 1685 --- DETKALPLPEASVKTDTGTESKPOGVIRS-POGLELALPSRDS- 2385 :|:| : : : : : | | ::: : | | XLSAVA------DDSLAVSHKDSLEASP--VLEDNSSHKTP 2419 CKEPKQEESSSSDPDADCSVDVDEPKHTGSGEDESGV-----PVL 2619 ARTPAKNLAPHHASPDPPAPPASASDPHREKTQSKPFSIQEL---- 2036 KHQPVSTAGKTEKHLPVSPSGKTEKQP-PVSPTSKTERIEETMSVRE 1959 VPYSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPP 1181 FATAMAORADMLRGLSPRESSLALNY----AAGPRGIIDLSQVPHLP 1741 ADRIAYLPTAPOPFSSRHSSSPLSP---GG--PTHLTKPTTTSSSER 1796 ORDREREKSILTSTTTVEHAPIWRPGTEQSSGSSGSSGGGGGSSSRP 1851 PRIQUALOORPSVLHNTG-------1891 PSKPTVLRSTSTSSPVRPAAT-FPPATHCPLGGTLDGVYPTLMEPVL 1937 KRPE----RPRADTGHAFLAKP-----PARSGLEPASSPSKGSE 1980 : || : ||-| : ||-| | sBPELAQLKKGADSG----E 2673

```
3290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | :: | | :: | | 3005
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----EIFNMPAITGTGLMTYRSQAVQEHASTNMGLEAIIRKALMGKYDQWEESPPL 2363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3066 NPFLFQEGKLFEMTR-SGAIDMTKRSYADESFHFFQIGQES------REETLSEDVKE 3116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2364 SANAFNPLNASASLPA------AMPITAAD----GRSDHTLTSPGGGGKA 2403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3117 GATGADPLPISAESLALSESKETVDDEADLLPDSVSEEVEETPASDAQLNSQMG---I 3173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2404 KVSGRPSSRKAKS----PAPGLASGDRPP-----SVSSVHSEGDCNRRTPLTN 2447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3174 SASTETPTKEAVSVGTKDLPTVQTGDIPPLSGVKQISCPDSSEPAVQVQLDF---STLTR 3230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2448 RVWEDR----PSSA-----GSTPFPYNPLIMRLOAGVMASPPPFGLPAGS 2488
                                                                                                                                                                                                                                                                                       VKGHQRVVTLAQHISEVITQDYTRHHPQQLSAPLPAPLYSFPGASCPVLDLRRPPSDLYL 2185
                                                                  DOPKICDGHGC----EAMSPSSSARPVSSGLQSPTGDDVDEQPVIYKESLALQGTHEKDT 2789
-PSL----THDKGL 2070
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EG--EELDVSRAES---POADCPSESFSSSSLPHCLVSEGKELDEDISATSSIOKTE--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2889 PQITSPYENVPSQSFFSSEESKTQIDANHTTSFHSSEVYSVTITSPVEDVVVASSSSGTV
                                                                                                                                         PKHLEELDKSHLEGELRPKQPGPVKLGGEAAHLPHL-----RPLPESQPSSSPLLQTAPG
                                                                                                                                                                                                                                                                                                                                                          -----VTKTDETFENLPKDCP---SQDSSITTQTDRFSM---DVPVSDLAE-NDEIYD
                                                                                                                                                                                                                                                                                                                                                                                                                                P---PPDHGAPARGSPHSEGGKRSPEPNKTSVLGGGE-DGIEPVSPPEGMTEPGHSRSAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ----PFSKLTESNSAM---VKSKKQEI-----NKKLNTHNRNEPEYNISQPGT----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDIJINE=94327588; PubMed=8051117;
Christiano A.M., Greenspan D.S., Lee S., Uitto J.;
"Cloning of human type VII collagen. Complete primary sequence of the
alpha 1(VII) chain and identification of intragenic polymorphisms.";
J. Biol. Chem. 269:20256-20262(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Buteleostomi,
Mammalia, Butheria, Primates, Catarrhini, Hominidae, Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE OF 128-1493 FROM N.A., AND PARTIAL SEQUENCE.
MEDLINE=93338437; PubMed=1307247;
Christiano A.M., Rosenbaum L.M., Chung-Honet L.C., Parente M.G.,
Woodley D.T., Pan T.C., Zhang R.Z., Chu M.-L., Burgeson R.E.,
Uitto J.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         collagen)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2242 YPLLYRDG-----EQTEPS--RMGSKSPGNTSQPPA---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CA17 HUMAN STANDARD; PRT; 2944 AA. 002388; Q14054; Q16507; Created) 01-UNN-1994 (Rel. 29, Created) 01-FRB-1996 (Rel. 33, Last sequence update) 10-OCT-2003 (Rel. 42, Last annotation update) Collagen alpha 1(VII) chain precursor (Long-Chain
   -ELRSLGYHGSSYSPEGVEPVS---PVSS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SVSEDFLSSVDEENK 3305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2489 GPLAGPHHAWDEEPK 2503
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Homo sapiens (Human).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  collagen).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3291
                                                                                                                                                                                                                                                                                                                                                                 2843
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2949
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2312
                                                                          2734
                                                                                                                                                                                                                                                                                                                                                                                                                                           2186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1D
DTT ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE ACCOUNT OF THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                      셤
                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                          Ś
                                                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ∂
                                                                                                                                                 ò
```

```
human type VII collagen gene (COL7A1),
previously characterized gene.";
                                                                                                                                                                                        the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Tanaka T., Takahashi K., Furukawa F., Imamura S.; "Molecular cloning and characterization of type VII collagen cDNA.", Biochem. Biophys. Res. Commun. 183:958-963(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               non-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE-9804169; PubMed-9375848;
Jaervikallio A., Pulkkinen L., Uitto J.;
Jandscular basis of dystrophic epidermolysis bullosa: mutations in the type VII collagen gene (COL7A1).";
Hum. Mutat. 10:338-347(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE=93291877; PubMed=8513326; Christiano A.M., Greenspan D.S., Hoffman G.G., Zhang X., Tamai Y., Christiano A.M., Greenspan D.S., Hoffman G.G., Zhang X., Tamai Y., Lin A.N., Dietz H.C., Hovnanian A., Uitto J.; Amissense mutation in type VII collagen in two affected siblings with recessive dystrophic epidermolysis bullosa."; Nat. Genet. 4:62-66(1993).
                                                                                                                                                                                                                                                                                                Gammon W.R., Abernethy M.L., Padilla K.M., Prisayanh P.S., Cook M.E., Wright J., Briggaman R.A., Hunt S.W. III; Moncollagenous (NCI) domain of collagen VII resembles multidomain adhesion proteins involved in tissue-specific organization of extracelular matrix."

J. Invest. Dermatol. 99:691-696 (1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE=94375010; PubMed=8088784;
Christiano A.M., Hoffman G.G., Chung-Honet L.C., Lee S., Cheng W.,
                                                                                                        SEQUENCE OF 815-1439 FROM N.A.
MEDLINE=91334380; PubMed=1871109;
MedlinE=91334380; PubMed=1871109;
Barente M.G., Chung L.C., Kyrnaenen J., Woodley D.T., Wyrn K.W.,
Bauer E.A., Mattei M.-G., Chu M.-L., Uitto J.;
"Human type VII collagen: cDNA cloning and chromosomal mapping of
           amino-terminal and chimeric. Homology to cartilage matrix protein, the type III domains of fibronectin and the A domains of von Willebrand factor.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Fine J.D., Johnson L., Wright T.;
"Epidermolysis bullosa simplex superficialis. A new variant of
epidermolysis bullosa characterized by subcorneal skin cleavage
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             the
The large non-collagenous domain (NC-1) of type VII collagen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         "The carboxyl-terminal half of type VII collagen, including collagenous NC-2 domain and intron/exon organization of the corresponding region of the COL7Al gene."; Hum. Mol. Genet. 2:273-278(1993).
                                                                                                                                                                                                                          Proc. Natl. Acad. Sci. U.S.A. 88:6931-6935(1991)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 [11]
WEALANT DEB SER-2040.
WEDLINE=94224777; PubMed=8170945;
Christiano A.M., Ryynaenen M., Uitto J.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              VARIANT EBDSC ARG-2034.
MEDLINE=89227237; PubMed=2653224;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       mimicking peeling skin syndrome.",
Arch. Dermatol. 125:633-638(1989)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE=92231902; PubMed=1567409;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE OF 2395-2944 FROM N.A. MEDLINE=93271985; PubMed=8499916;
                                                                                                                                                                                                                                                                                       MEDLINE=93107742; PubMed=1469284;
                                                                          Genet. 1:475-481(1992)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Uitto J., Greenspan D.S.;
"Structural organization of the
composed of more exons than any
Genomics 21:169-179(1994).
                                                                                                                                                                                                                                                                   SEQUENCE OF 369-1255 FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE OF 340-675 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE OF 1-87 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REVIEW ON DEB VARIANTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                            TISSUE=Keratinocytes;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                [10]
VARIANT DEB LYS-2798.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               rissum=Placenta;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Greenspan D.S.;
                                                                                                                                                                                                                gene.";
```

DEB ARG-2009 AND ARG-2043

```
MEDLINE=98106792; PubMed=9444387; Cserhalmi-Friedman P.B., Karpati S., Horvath A., Christiano A.M.; Cserhalmi-Friedman P.B., Karpati S., Horvath A., Christiano A.M.; "Identification of a glycine substitution and a splice site mutation in the type VI collagen gene in a proband with mitis recessive dystrophic epidermolysis bullosa."; Arch. Dermatol. Res. 289:640-645(1997).
                                                                                                                                                                                                                                                                                     MEDLINE-95164985; PubMed=7861014; Christiano A.M., Morricone A., Paradisi M., Angelo C., Mazzanti C., Cavalieri R., Uitto J.; Cavalieri R., Uitto J.; Maycine-to-arginine substitution in the triple-helical domain of type VII collagen in a family with dominant dystrophic epidermolysis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Christiano A.M., McGrath J.A., Uitto J.;
"Influence of the second COL7A1 mutation in determining the
phenotypic severity of recessive dystrophic epidermolysis bullosa.";
J. Invest. Dermatol. 106:766-770(1996).
                                                                                                                    MEDLINE=96081220; PubMed=8541842;
Christiano A.M., Lee J.Y.-Y., Chen W.J., Laforgia S., Uitto J.,
"Pretibial epidermolysis bullosa: genetic linkage to COL7A1 and
identification of a glycine-to-cysteine substitution in the triple-
helical domain of type VII collagen.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Shimizu H., McGrath J.A., Christiano A.M., Nishikawa T., Uitto J., "Molecular basis of recessive dystrophic epidermolysis bullosa: genotype/phenotype correlation in a case of moderate clinical
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE=96220218; PubMed=8644729;
Christiano A.M., McGrath J.A., Tan K.C., Uitto J.;
"Glycine substitutions in the triple-helical region of type VII
collagen result in a spectrum of dystrophic epidermolysis bullosa
phenotypes and patterns of inheritance.";
Am. J. Hum. Genet. 58:671-681(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             VARIANTS DEB TRP-1982; GLY-2008; ALA-2025; GLU-2049; TRP-2063 AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hovnanian A., Rochat A., Bodemer C., Petit E., Rivers C.A., Prosteritag S., Christiano A.M., Uitto J., Lathrop M., Barrandon Y.,
  "Dominant dystrophic epidermolysis bullosa: identification of a
31y-->Ser substitution in the triple-helical domain of type VII
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   "Characterization of 18 new mutations in COL7A1 in recessive dystrophic epidermolysis bullosa provides evidence for distinct molecular mechanisms underlying defective anchoring fibril
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Dunnill M.G.S., McGrath J.A., Richards A.J., Christiano A.M., Uitto J., Pope F.M., Eady R.A.J.;
"Clinicopathological correlations of compound heterozygous C'annicopathological correlations of compound in recessive dystrophic epidermolysis bullosa.";
J. Invest. Dermatol. 107:171-177(1996).
                                                            U.S.A. 91:3549-3553(1994)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Invest. Dermatol. 106:119-124(1996)
                                                                                                                                                                                                                                                                                                                                                                                                                    Invest. Dermatol. 104:438-440(1995)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       J. Hum. Genet. 61:599-610(1997)
                                                                                                                                                                                                                             Mol. Genet. 4:1579-1583(1995)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE=96310789; PubMed=8757758;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE=97465605; PubMed=9326325;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE=96154068; PubMed=8592061;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE=96183562; PubMed=8618018;
                                                            Sci.
                                                                                                                                                                                                                                                                    VARIANT DEB ARG-2043.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VARIANT DEB ASP-2073.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VARIANT DEB ARG-1652.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VARIANT DEB ARG-2575.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VARIANT DEB ARG-1782
                                                            Proc. Natl. Acad.
                                                                                                                                                                                                                                                                                                                                                                                                                                                           VARIANTS DEB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Prost Y.;
                                                                                                                                                                                                                                                                                                                                                                           type VII c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      severity.
                                        collagen
                                                                                                                                                                                                                                 Hum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          191
```

```
1537 GRP--GDPAVVGPAVAGPKGEKGDVGP-AGPRGATGVQGERGPPGLVLP-GDPGPKGDPG 1592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                958 - QKPLDLKQLKQRAAAIPPIQVTKVHEPPREDAAPTKPAPPAPPPPQNLQPESDAPQQPG 1016
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1017 SSPRGKSRSPAPPADKEAFAAEAQKLPGDPPCWTSGLPFPVPPREVIKASPHAPDPSAFS 1076
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1627 VGPRGRDGEVGEKGDEGP-----PGDP-----GLPGKAGERG-LRGAPGVRGPVGEK 1672
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -----YAPPGHPLPLGLHDTARPVLPRPFTISNPPPLISSAKHPSVLERQIGAI 1125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1126 SQGMSVQLHVPYSEHAKAPVGPV-TMGLPLPMDPKKLAPFSGVKQEQLSP--RGQAGPPE 1182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1303 ADGRPGSPGRAGNPGTPGAPGLKGSPGLPGPRGDPG-ERGPRGPKGEPGAPGQVIGGEGP 1361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1362 GLPGRKGDPGPSGPPGPRGPLGDPGPRGPPGLPGTAMKGDKGDRGERGPPGPGEGGIAPG 1421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1422 EPGLPGLPGSPGPPGPVGPPGKKGEKGDSEDGAPGLPGQPGSPGEQGPRGPPGAIGPKGD 1481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1183 SLGVPTAQEASVL---RGTALGSVPGGSITK-----GIPSTR----VPSDSAITYRGS 1228
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1771 DRGPPGLDGRSGLDGKPGAAGPSGPNGAAGKAGDPGRDGLPGLRGEQGLPGPSG----- 1824
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 854
                                                                                                      orysis bullosa by mutation in the
                                                                                                                                                                                                                                                                                                        result in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ---RTSRAPIEPTPASEATG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   726 ASGNEVPRGECSGPAT -- VNNSSDTESIPSPHTEAAKDTGONGPK----PPATLGADGP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 805 APTPPPAPPSPSAPPPVVP---KEEKEEETAAAP-----PVEEGEE--QKPPAAEELAVD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TGKAEEPVKSECTEEAEEGP-----AKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----GRLVDTGPGAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1717 EKG-----EPGDRGOEGPRGPKGDPGLPGAPGERGIEGFRGPPGPQGDPGVRGPAGEKG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         910 GAPQDSDSSATCSADEVDEAEGGDKNRLLSPR------PSLLTPTGDPRANASP-
                                                                                                        dystrophic epidermolysis bullosa
                                           Anton-Lamprecht I.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1593 DRGPIGL----TGRAGPPGD----SGPPGEKGDPGRPGPPGP-----
                                                                                                                                                                                                                               MEDLINE=98334662; PubMed=9668111; Hammani-Hauasli N., Schumann H., Raghunath M., Kilgus O., Luethi Luger T., Bruckner-Tuderman L.; Rome, but not all, glycine substitution mutations in COL7A1 resuintracellular accumulation of collagen VII, loss of anchoring
                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE=98410969; PubMed=9740253;
Kon A., Pulkkinen L., Ishida-Yamamoto A., Hashimoto I., Uitto
"Novel COL7A1 mutations in dystrophic forms of epidermolysis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3.0%; Score 390; DB 1; Length 2944; llarity 22.0%; Pred. No. 1.3e-05; Conservative 141; Mismatches 773; Indels 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Terracina M., Posteraro P., Schubert M., Sonego G., Atzori
                                                                                                                                                                                                               VARIANTS DEB ASP-1519; ASP-2006; GLU-2015 AND ARG-2034
               MEDLINE=97358588; PubMed=9215684; Wilbsen O., Anton-L. Winberg J.-O., Hammami-Hauasli N., Nilssen O., Anton-L. Naylor S.L., Kerbacher K., Zimmermann M., Krajol P., T., Bruckner-Tuderman L.; Modulation of disease severity of dystrophic epidermo a splice site mutation in combination with a missense
                                                                                                                                                                                                                                                                                                                                                                                                                 VARIANTS DEB CYS-2008; ARG-2207 AND SER-2775.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1673 GDQGDPGEDGRNGSPGSSGPKG--DRGEPGPPGPP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Invest. Dermatol. 111:534-537(1998)
                                                                                                                                                                                                                                                                                                                                               fibrils, and skin blistering.";
J. Biol. Chem. 273:19228-19234(1998)
                                                                                                                                                                      Mol. Genet. 6:1125-1135(1997)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE=99019477; PubMed=9804332;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VARIANT DEB ARG-1347.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches 434;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      bullosa."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       855
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          779
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1077
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   23]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                بر
ن:
셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              · 상 · 원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

us-09-522-753-5.rsp

2189 DHGA 2669 QSGA 2228 2729 SVGP 2275 LITES 2324 LMTY 2334 LMTY 2336 2382 PI 2382 PI	RESULT 36	- 1 4 4 8 8 8 4 7 1 日日 2 7
86868686	######################################	8888888888888
1229 ITHGTPADVLYKGTITRIIGEDSPSKLDRGREDSLPKGHVIYEGKKCHVLGYEGGMSVTQ 1288		STATEMENT OF THE CONTROL OF THE CONT
8 6 8 6 8 6 8 6	*	3 6 6 6 6 6

```
PROT entry is copyright. It is produced through a collaboration e Swiss Institute of Bioinformatics and the EMBL outstation no Bioinformatics Institute. There are no restrictions on its n-profit institutions as long as its content is in no way it that statement is not removed. Usage by and for commercial quires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                         -VSHAEEE------ERVPPEDDEYSEYSEXSUEXQDPEAPWDSDDPCSL 2878
                                                                                                                                ---PEGMT------EPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQPPAFFSK 2274
                                                                                                                                                                                                PRGPEGLOGOKGERGPPGERVVGAPGVPGAPGERGEOGRPGPAGPRGEKGEAA---- 2784
APAR-----GEDG----GSPHSEGGKRSPEPNKTSVLGG-----GEDG-IEPVSP-- 2227
                                                                   APGKEGLIGPKGDRGFDĠQPGPKGDQGEKGERGTPGIĠĠFPGPSĠNDĠSAGPPGPPG 2728
                                                                                                                                                                                                                                                                          SNSAMVKSKKQEINKKLNTH------NRNEPEYNISQPGTEIFNMPAITGTG 2323
                                                                                                                                                                                                                                                                                                                                                                                                              YRSQAVQEHASTNMGLEAIIRKALMGKYDQWEESPPLSANAFNPLNA--SASLPAAM 2381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Acad. Sci. U.S.A. 90:6350-6354 (1993). We Possibly acts as a transcriptional regulatory factor. Intee genes involved in skeletal formation during embryonic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -----TAADGRSD--HTLTSPGGGGKAKVSGRPSSRKAKSP 2417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | | : : | | : : | | : : | EGSCTAYTLRWYHRAVTGSTEACHPFVYGGCGGNANRFGTREACERRCP 2930
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NM N.A. (ISOFORMS 1 AND 2).

1/6J, and CS7BL/6 X CBA; TISSUE=Lung, and Spleen;

1/76J; PubMed=8375517,

CTGC C.M., Siracusa L.D., Buchberg A.M.;

CTCC C.M., Siracusa L.D., Buchberg A.M.;

The murine All-1 gene reveals conserved domains with and identifies a motif shared with DNA carages.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    letazoa, Chordata, Craniata, Vertebrata, Euteleostomi, theria, Rodentia, Sciurognathi, Muridae, Murinae, Mus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Interacts with SBF1 (By similarity).
JLAR LOCATION: Nuclear (By similarity)
FIVE PRODUCTS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          =P55200-2; Sequence=VSP 006667;
ITY: Belongs to the TRX/MLL family.
ITY: Contains 1 bromodomain.
ITY: Contains 1 SET domain.
ITY: Contains 3 PHD-type zinc fingers.
ITY: Contains 1 CXXC-type zinc finger.
ITY: Contains 1 post-SET domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ternative splicing; Named isoforms=2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (Rel. 34, Last sequence update)
(Rel. 41, Last annotation update)
protein HRX (ALL-1) (Fragment).
OR ALL1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT; 3866 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 P55200-1; Sequence=Displayed;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    smail to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MGD; MGI:96995; Mll.
GO; GO:0005634; C:nucleus; IDA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (Rel. 34, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9; AAA62593.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (Mouse)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  10600
```

```
1059 QVPEDCGICTNCLDKPKFGGRNIKKQCCKMRKCQNLQWMPSKASLQKQTKAVKKKEKKSK 1118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1179 APEPKQVSAPASRKSSKQ------1196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----VSQPAAVVPOPPORTAPOKKEAPKAVP-----SEPKKK-----QPPPPE 1235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PPPPQNLQ-----PESDAPQQPGSSPRGKSRSPAPPADKEAFA------AEA 1039
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PGPEQSKQKKVAPLPSIPVKQKP-----KDKEKPPPVSKQENAGTLNILNPLSNGISSK 1289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1040 QKLPGD------PPC----WTSG-----LPFPVPPREV---IKASPHAPDPSAFS 1076
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Y----APPGHPLPLGLHDTARPVLPRPFTISNPPPLISSAKHPSVLERQIGAISQGMSVQ 1132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1346 YCQVCCEPFH--KFCLEENERP------1380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TAQE----ASVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYKGT 1242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   .017 EDAEPLAPP-----CGQCPGC 1058
                                                                                                     ---ASALNPIFIFPSHS 700
                                                                                                                                                                                                                                                                        LYPVGRVSKEK---VAGEDVGTSSSAKKATGRKKSSSLDSGADVAPVTLGDTTAVKAKIL 858
                                                                                                                                                                                                                                                                                                                           KTVSQCKNFYFNYKKRQNLDEILQQHKLKMEKERNARRK-------KKK 685
                                                                                                                                                                                                                                                                                                                                                                                   APAAASEEAAFPPVVEDEEMEASGVSGNEEEMVEEAEALHASGNEVPRGECSGPATVNNS 745
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SDTESIPSPHTEAAKDTGQNGPKPPATLGADGPPPGPPTPPRTSRAPIEPTPASEATGA 805
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SPRPSLLTPTGDPRANASPQKPLDLKQLKQRAAAIPPIQVTKVHEPPREDAAPTKPAPPA 998
                  KKVERIENNPRRRAKESKV--REYYEKQFPEIRKQ-RELQERMQSRVGQRGSGLSMSAAR 367
                                                                        SEHEVSEIIDGLSEQENLEKQMRQLAVIPPMLYDADQQRIKFINMNGLMADPMKVYKDRQ 427
                                                                                                                                                                                                           ITP-QQSAELASMELNESSRWTEEE----METAKKGLLEHGRN--WSAIAR-----MVGS 643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VMNMWSEQEKETFREKFMQHPKNFGLIASFLERKTVAECVLYYYLTKKNENYKSLVRRSY
                                                                                                                                                   -----DLKEKTDDTSGED-NDEKEAVASKGRKTANS--QGRRKGRITRSMANEANSEEA
                                                                                                                                                                                                                                                                                                                                                                                                                                                       AVALGRKRAVFP----DDMPTLSALPWEEREKI-----LSSMGND-----DKSSVAGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----EETAAAPPVEEGEEQKPPAAEELAVDTGKAEEPVKSECTEEAEEG--PAKGK
                                           ESVILPSNRISSGASSSGVSNRKRKRKKVFSPIRSEPRSPSHSMRIRSGR----LSTS---
                                                                                                                                                                                         RRRGKSQQQQQQQQQQQQQQQPMPRSSQEEKDEKEKEAEKEEEKPEVENDKED---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DABAAEATAEGALKAEKKEGGSGRATTAKSSGAPQDSDSSATCSADEVDEAEGGDKNRLL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                OKIPADGVHRIRVDFKEDCEAENVWEMGGLGILTSVPITPRVVCFLCSSSEHV----EFV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LHVPYSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLG----VP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1381 CHVCGROHOAT------KOLLECNKCRNSY-----HPECLGPNYPTKP
                                                                                                   ---ELSPLTPPSSVSSSLSIPVSPLA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1077
                                                                                                                                                                                                                      745
                                                                                                                                                                                                                                                                               802
                                                                                                                                                                                                                                                                                                          969
                                                                                                                                                                                                                                                                                                                                       859
                                                                                                                                                                                                                                                                                                                                                                    644
                                                                                                                                                                                                                                                                                                                                                                                               919
                                                                                                                                                                                                                                                                                                                                                                                                                           989
                                                                                                                                                                                                                                                                                                                                                                                                                                                        973
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     746
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     939
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1197
                                             611
                                                                                                     664
                                                                                                                                 428
                                                                                                                                                              701
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            879
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               666
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1188
                                                                                                                                                                                                                                                  545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    828
                  ŝ
                                       쉽
                                                                        ઠે
                                                                                               셤
                                                                                                                              ò
                                                                                                                                                        셤
                                                                                                                                                                                         ò
                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                 요
                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                  요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          610
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ----GKSEMEF---IESKRPRLELLPD------PLLRPSPLLATGQPA-GSEDLTKD 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RSLTGKLEP-VSPPSPPHTDPELEL-VPPRLSKEELIQNMDRVDREITMVEQQISKLKKK 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SKRYSMSERSFGSRATKKLPTLQSAPQQKTSSSPPPPLFTPPPPL---QPASG1SDHT-- 480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STQLVAQTWRATEPR-YPPHSLSYPVQIAR-----THTDVGLLEYQHHSRDYASHL 53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | | : ||: : :: |:|: || SSDS---SRSSSPSI--DTTSDSQASEEIQALPEERSNTPEVHTPLPISQSPENESNDRR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -----PWLMPPTIPLASPFLPASAAPMOGKRKSIL-----REPTFRWTSLKHSRSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        579 ---PLH---SGTRFDIHKRSPILRAPR----FTPSEAHSR------IF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SPGSIIQPQRRRPSLLSEFQPGNERSQELHLRPESHSYLPEL--------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QQQLEE-EAAKPPEPEKPV----SPPPIESKHRSLVQIIYDENRKKAEAAHRILEGLGPQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                527 PQYFSSAKYAKEGLIRKPIFDNFRPPPLTPED---VGFASGFSASGTAASARLFS----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             VELPLYNQPSDTR-QYHENIKINQAMRKKLILYFKRRNHARKQWKQKFCQRYDQLMEALE
                                                   Query Match 2.9%; Score 389.5; DB 1; Length 3866; Best Local Similarity 18.3%; Pred. No. 1.8e-05; Matches 591; Conservative 332; Mismatches 1081; Indels 1225;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3866 AA; 420976 MW; ADFC55E14E806F1D CRC64;
                                                                                                                                                                                                                                                                                                                                                                              SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BROMODOMAIN (DIVERGENT)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (in isoform 2)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POLY-GLU.
POLY-GLU.
Missing (in isoform
/FTId=VSP_006667.
                                                                                                                                                                                                                                                                                                                                                                             A.T HOOK (BY S
A.T HOOK (BY S
A.T HOOK (BY S
CXXC-TYPE.
PHD-TYPE 1.
PHD-TYPE 2.
PHD-TYPE 3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POST-SET.
POLY-GLY.
PRO-RICH.
POLY-PRO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 F-
^-
 Bromodomain.
FYrich C.
FYrich N.
PostSET.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           469
1238
3536
3697
1505
                                                                                                                                                                                                                                                                                                                                                                                               125
207
207
1091
1381
1432
1529
1650
3846
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1497
InterPro; IPR001487;
InterPro; IPR003889;
InterPro; IPR003888;
                                           InterPro; IPR003616;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3533
3693
1503
                                                                                                                                                                                                                                                                                                                                                Polymorphism.

NON TER

NON TER

DNA_BIND 11

DNA_BIND 15

ZN_FING 104

ZN_FING 13

ZN_FING 13

ZN_FING 13

ZN_FING 146

DOMAIN 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                316
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
```

a

ઠે 8 ઠે 요

ઠે 셤 ò 셤 ò

```
recessive Alport syndrome.";
                                                                                                                                                                                                                                                                                                      +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -
 Heidet L., Arrondel C., Forestier L., Cohen-Solal L., Mollet G., Gutierrez B., Stavrou C., Gubler M.C., Antignac C.; "Structure of the human type IV collagen gene COL4A3 and mutations in
                                                                                                                                                                                           Quinones S., Bernal D., Garcia-Sogo M., Elena S.F., Saus J.; Quinones S., Bernal D., Garcia-Sogo M., Elena S.F., Saus J.; Goodpasture attigen (alpha 3(V)NCI). Identification of a potentially antigenic region at the triple hellx/NCI domain junction."; J. Biol. Chem. 267:19780-19784(1992).
                                                                                                                                                                                                                                                                                       MEDLINE=91353570; PubMed=1882840;
Morrison K.E., Mariyama M., Yang-Peng T.L., Reeders S.T.;
"Sequence and localization of a partial cDNA encoding the human alpha
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   "Mutations in the type IV collagen alpha 3 (COL4A3) gene in autosomal
                        SEQUENCE FROM N.A., VARIANTS AS GLU-297; ARG-407; ARG-640; ARG-1167; GLU-1207; GLU-1207; GLU-1207; GLU-137; FRR-1277; THR-1330; GLU-1334; GLU-1347 AND CYS-1661, AND VARIANTS ARG-43; GLU-162; TYR-326; HIS-408; ARG-451; PRO-574; GLU-1269 AND PRO-1474.
                                                                                                                                                                                                                                                                                                                                                                                                                      Pusey C.D.; "Molecular cloning of the human Goodpasture antigen demonstrates it to be the alpha 3 chain of type IV collagen."; J. Clin. Invest. 89:592-601(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 VARIANT PRO-1474.
WEDLINE=55078877. PubMed=7987301;
Lemmink H.H., Mochizuki T., van den Heuvel L.P.W.J., Schroeder C.H.
Barrientos A., Monnens L.A.H., van Oost B.A., Bruner H.G.,
Reeders S.T., Smeets H.J.M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         "Two genes, COL4A3 and COL4A4 coding for the human alpha3(IV) and alpha4(IV) collagen chains are arranged head-to-head on chromosome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Bernal D., Quinones S., Saus J.;
"The human mRNA encoding the Goodpasture antigen is alternatively
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE=94124597; PubMed=8294492;
Feng L., Xia Y., Wilson C.B.;
"Alternative splicing of the NC1 domain of the human alpha 3(IV)
collagen gene. Differential expression of mRNA transcripts that
predict three protein variants with distinct carboxyl regions.";
J. Biol. Chem. 269:2342-2348(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Momota R., Sugimoto M., Oohashi T., Kigasawa K., Yoshioka H., Ninomiya Y.;
                                                                                                                                                                                                                                                                                                                                                                                                          Turner N., Mason P.J., Brown R., Fox M., Povey S., Rees A.,
Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ding J.;
Submitted (JAN-1993) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                     SEQUENCE OF 1386-1670 FROM N.A., AND PARTIAL SEQUENCE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE OF 1439-1670, AND ALTERNATIVE SPLICING
                                                                                                                                           Am. Soc. Nephrol. 12:97-106(2001).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Biol. Chem. 268:12090-12094(1993)
                                                                                                                                                                                                                                                                                                                               3 chain of type IV collagen.";
Am. J. Hum. Genet. 49:545-554(1991).
                                                                          MEDLINE=21064696; PubMed=11134255;
                                                                                                                                                                                   LINE=93015826; PubMed=1400291;
                                                                                                                                                                                                                                                                                                                                                                              TISSUB-Kidney;
MEDLINE-92147878; PubMed=1737849;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE=98196854; PubMed=9537506;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE=93280184; PubMed=8505332;
                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE OF 1331-1670 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE OF 1644-1670 FROM N.A.
                                                                                                                                                                                                                                                                           SEQUENCE OF 1453-1670 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FEBS Lett. 424:11-16(1998).
                                                                                                                                 autosomal Alport syndrome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE OF 1-29 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ALTERNATIVE SPLICING.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TISSUE=Kidney;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TISSUE=Kidney;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                spliced.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5
```

```
rance-3; syloupus-13; Sequence=VSP 001171; collagens are colocalized and present only in basement membranes of kidney, eye, colocalized and present only in basement membranes of kidney, eye, cochlea, lung and brain.

-!- DOMAIN: Alpha chains of type IV collagen have a noncollagenous domain (NCI) at their C-terminus, frequent interruptions of the G-X-Y repeats in the long central triple-helical domain (which may cause flexibility in the triple helical domain (which may cause flexibility in the triple helix), and a short N-terminal triple-helical 7S domain.

-!- FTM: Prolines at the third position of the tripeptide repeating unit (G-X-Y) are hydroxylated in some or all of the chains.

-!- FTM: Type IV collagens contain numerous cysteine residues which are involved in inter- and intramolecular disulfide bonding. 12 of these, located in the NCI domain, are conserved in all known type IV. collagens.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PTM: Phosphorylated by the Goodpasture antigen-binding protein. DISEASE: Antibodies against the NC1 domain of alpha3 (IV) mediate the autoimmune disease Goodpasture syndrome [MIM:23450], which is characterized by hematuria and pulmonary hemorrhage. DISEASE: Defects in COL4A3 are a cause of autosomal recessive Alport syndrome (AS) [MIM:203780], an hereditary disorder characterized by progressive glomerulonephritis, renal failure, hematuria, ocular abnormalities and deafness. The recessive form occurs equally between males and females.
                                                                                                              nidogen.
SUBUNIT: There are six type IV collagen isoforms, alpha 1(IV)-
alpha 6(IV), each of which can form a triple helix structure with
2 other chains to generate type IV collagen network.
SUBGELIULAR LOCATION: Cell surface (Potential).
ALTERNATIVE PRODUCTS:
                                                                                                                                                                                                                                                                     Event=Alternative splicing, Named isoforms=3;
Comment=Additional isoforms seem to exist. Isoforms differ in
the C-terminal part of the NC1 domain;
                                    ö
Hum. Mol. Genet. 3:1269-1273(1994).
-!- FUNCTION: Type IV collagen is the major structural component glomerular basement membranes (GBM), forming a 'chicken-wire' meshwork together with laminins, proteoglycans and entactin/
                                                                                                                                                                                                                                                                                                                                                                                                                      Name=2; Synonyms=V;
IsoId=Q01955-2; Sequence=VSP_001170;
                                                                                                                                                                                                                                                                                                                                                                                             IsoId=Q01955-1; Sequence=Displayed;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           JOINED.
JOINED.
JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   JOINED.
JOINED.
JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AJ288487; CAC36101.1; -. AJ288488; CAC36101.1; JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CAC36101.1;
CAC36101.1;
CAC36101.1;
CAC36101.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CAC36101.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CAC36101.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AJ288489; CAC36101.1
AJ288490; CAC36101.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; X80031; CAA56335.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CAC36101.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CAC36101.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CAC36101.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CAC36101.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CAC36101.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Name=3; Synonyms=L5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AJ288493;
AJ288494;
AJ288495;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AJ288503;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AJ288496;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AJ288499;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AJ288500;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AJ288491
                                                                                                                                                                                                                                                                                                                                                                   Name=1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL;
                                                                                                                                                                                                                        ++
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 --
                                                                                                                                           -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        <del>:</del>
```

CAC36101.1; CAC36101.1;

AJ288504; (AJ288505; (

```
2235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  --LLYRDGEQ----TEPSRMGSKSPGNTSOPPAFFSKLTE 2277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2011 PPAPPASASDPHREKTQSKPFSIQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDKGL 2070
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ----GFQGFPGVKGEKGNPGFLG-SIGPPG--PIGPKGPPGVRGDPGT 1350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1892
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRADIGHAFLAKPPARSGLEPASSPSKGSEPRPLVPPVSGHATIARTPAKNLAPHHASPD 2010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----PGMOGE- 1374
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QGAKGDRGAPGFPGLPGKKGAMGDA-----GPRGPTGIEGFP-----GPPGLPGAIIP 1235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SKPTVLRST--SISSPVRPAATFPPATHCPLGGTLDGVYPTLMEPVLLPKEAPRVARPER 1950
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----APG 1298
                                                                                                 1409
                                                                                                                                                                                                                                                                                  1522 JARGAPVIVPE---LGKPROSPLTY--EDHGAP----FAGH-----LP----R 1556
                                                                                                                                                                                                                                                                                                                                                                          GLPGPAGPPGPR----GDLGSTGNPGEPGLRGIPGSMGNMGMPGSKGKRGTLGFPGRAGRP 1041
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PHLYPPYLIRGYPDTAALENRQTIINDYITSQQMHHNTATAMAQRADMLRGLSPRESSLA 1721
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LNYAAGPRGIIDLSQVPHLPVLVPP----TPGTPATAMDRLAYLPTAPQPFSSRHSSSPL 1777
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1410 AHEGLVATVKEAGRSIHEIPREELRHTPELP--LAPRPLKEGSITOGTPLKYDTGASTTG 1467
                                                                                                                                                                                                                       1468 SKKHDVRSLIGSPGRTFPPVHPLDVMADARALERACYEESLKSRPGTASSSGG-----S 1521
                                                                                                                                                                                                                                                      924
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PKHLEELDKSHLEGELRPKQPGPVKLGGEAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GAPAR-----GSPHSEGGKRSPEPNKTSVLGG-----GEDGIEPVSPPEGMTEPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RVVTLAQHISEVITQDYTRHHPQQLSAPLPAPLYSFPGASCPVLDLRRPPSDLYLPPPDH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | | : | : | | | | | GKPGKRGDSGSPATWTTRGFVFTR
                                                                                                                                                                                                                                                                                                      925 GQRGTPGAKGEQGDKGNPGPSEISHVIGDKGEPGLKGFAGNPGEKGNRGVPGMPGLKGLK
                                                                                                                                                                                                                                                                                                                                                                                                         1608 ISPYEHLLRGVSGVDLYRSHI-PLAFDPTSIPRGIPLDAAAAVYL----PRHLAPNPTY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SGSSGGGG-----GSSSRPASHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGIITAVBP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GQ-TGNKGPPGSRGSPGAPGPPGPPGSHV-IG--IKGDKGSMGHPG--PKGPPGTA----
                                                                                                                                                                                                                                                       ----GNPGTPGQRGSPGIPGVK
                                                                                                                                                                                                                                                                                                                                              GSPVTMREPTPRLQEGSLSSSKASQDRKLTSTPREI -----AKSPHSTV--PEHHPHP
                                                                                                                                                                                                                                                                                                                                                                                                                                     1042 GLPGIHGLQGDKGEPGYSEGTRPGPPGPTGDP-GLPGDMGKKGEMGQPGPPGHL--GPAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SPGGPTHLTKPTTTSSSERERDRDRERDRDREREKSILTSTTTVEHAPIWRPGTEQSSGS
                                                                                               1354 SI----TQGIPRSYVEAQEDYLRREAKLLKREGTPPPPPPPSRDLTEAYKTQALGPLKLKP
                                    1294 GRSSSGPPHETAAPKRTYDMMEGRVGRAISSASIEGLMGRAIPPERHSPHHLKEQHHIRG
                                                                    -----LKGQQGRRG
                                                                                                                                 ----ETGSPG----IP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PEGAP----GSPGSPGLPGK-----------
                                                                                                                                                                                            862 GHQGEMGPLGQRG------YPGNPGILGP-PGEDGVI------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ----pgphgblgfkglkgl--Lgppglkgppglpg-----Fpgspgpmglk-
                                                                    ----PGPPGEQGPPGRCIEGPRGAQG-----LPGLNG-
                                                                                                                                 ---LDRSGFPG---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                --GDMG-----PPGRLG-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LKIIS------LPGSPGPPGE
           -------GLDGP-RGDPGQ
                                                                                                                                                                                                                                                          -----GMMGFPGAIGPPGPP-
                                                                                                                                   KTGPKGDPGİPG---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TPGLPGPRGDP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HSRSAVYP-
- :
PGNE--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1838
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2071
                                                                                                                                                                                                                                                                                                                                                      1557
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1662
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1099
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1722
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1893
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1375
                                                                                                                                                                                                                                                            892
                                                                                                                                    834
              782
                                                                         797
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8
                                                                                                                                 g
                                                                                                                                                                   8
                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                    유
                                                                      g
                                                                                                       ò
                                                                                                                                                                                                                               ਨੇ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1234 PADVLYKGTITRIIGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKED 1293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GAISQGMSVQLHVPYSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSP--RGQAGP 1180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1003 QNLQPESDAPQQPGSSPRGKSRSPAPPADKEAFAAEAQKLPGDPPCWTSGLPFPVPPREV 1062
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              781
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   731
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                448
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAEEL------AVDTGKAEEPVKSECTE-EAEEGPAKGKDAEAAEATAEGALKAEK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             797 -----TPASEATGAPTPPPAPPSPAP----PPVVPKEEKBEETAAAPPVEEGEEOKPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             402 SKGERGRÞGKDAMGTÞGSÞGCAGSÞGLÞGSPGÞÞ------GÞÞGDIVFRKGÞP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRANASP-QKPLD-----LKQLKQRAAAIPPIQVTKVHEPPREDAAPTKPAPPAPPPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PGLRGQPGRKGLDGIPGTLGVKGLPGPKGELALSGEKGDQGPPGDPGSPGSPGPAGPAGP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1063 IKASPHAPDPSAFSYAPPGHPLPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLERQI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LSVSTPVPGPPG-PPGPPGHPGPQG------PPGI---PGSLGKCGDPGL----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PESLGVPTAQEASVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSI-----THGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PGKPGLPGA-----KGEPAVAMPGGPGTPGFPGERGNSGE----HGEIGLPGLPGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PSPHTEAAKDTGQNG-PKPPATLGADGP----PPGPPTPPRRTSR-----APIEP---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     449 GDHGLPGYLGSPGIPGVDGPKGEPGLL--CTQCPYIPGP-----PGLPGLPGLHG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     896 KEGGSGRATTAKSSGAPQDSDSSATCSADEVDEAEG----GDKNRLLSPRPSLLTPTGD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -------PGYGPQG-----EPGLQGTQGVPGAP-----GPPGEAGPRGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ----PGPDGEPGIP------GIGFPGPPGPKGDQGFPGTKGSLGCPGKMGEPGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          740;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 1670;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2.9%; Score 389; DB 1; Length 16
larity 20.8%; Pred. No. 8.5e-06;
Conservative 119; Mismatches 610; Indels
   JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
                                                                                                                                                                                                                                                                 JOINED.
                                                                                                                                                                                                                                                                                               JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CAC36101.1;
                                                                                                                            CAC36101.
                                                                                                                                                                                                                                    CAC36101.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity
                                                                                 AJ288511;
AJ288512;
                                                                                                                                                                                                                                                                                    AJ288524;
                                                                                                                                                                                                                                                                                                 AJ288525;
                                                                                                               AJ288513;
                                                                                                                                                                                                                                       AJ288521;
                                                                                                                                                                                                                                                                      AJ288523;
                                                                                                                                              AJ288515;
                                                                                                                                                                                                                                                                                                                                              AJ288528;
                                                                                                                                                                            AJ288517
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               386;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        951
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               648
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   732
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             752
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          343
                                                                                                                                                                                                                                                                                               EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                               EMBL;
EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL;
EMBL;
                                                   EMBL;
EMBL;
EMBL;
EMBL;
EMBL;
EMBL;
EMBL;
EMBL;
                                                                                                                                                                                       EMBL;
EMBL;
                                                                                                                                                                                                                      EMBL;
EMBL;
                                                                                                                                                                                                                                                    EMBL;
EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                          EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                            EMBL:
                                                                                                                                                                                                                                                                                                                                                                                                          EMBL;
```

8 8 8 8 8

원

6 B 6

8 6

셤

à

성 유

ò

64;

313 255

```
104 ESKRPRLELLPDPLLRPSPLLATGQPAGSEDLTKDRSLTGKLEPVSPPSPPHTDPEL--E 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       256 LYNQPSDTRQYHENIKINQAMRKKLILYFKRRNHARKQWK-------QKFCQ--R 301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ----PEKSSKVSEQLKCCSGILKEMP----AKKHAAYAWPFYKPVDVEALGLHDYCDIIK 395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       302 YDQLMEALEKKVERIENNPRRRAKE-----SKVREYYEKQFPEIRKQRELQERMQS 352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        396 HPMDMSTIKSKLEARE---YRDAQEFGADVRLMFSNCYKYNPPDHEVVAMARKLQDVFEM 452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      413 NGLMADPMKVYKDRQVMNMWSEQEKETFREKFMQHPKNFGLIASFLERKTVAECVLYYYL ·472
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           583 TRSMANBANSEBAITPQQSAE--LASMELNESSRWTEBEMETAKKGLLEHGRNWSAIARM 640
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BLAVDTGKAEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSS 909
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             --ILHLPQPELPPHLPQPPEHSTPPHLNQHAVVSPPALHNAL-PQQPSRPSNRAAALPPK 883
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         473 TKKNENYKSLVRRSYRRRGKSQQQQQQQQQQ----QQQQQQQPMPRSSQBEKDEKEKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           529 AEKEEEKPEVENDKEDLLKE-----KTDDTSGEDNDEKEAVASKGRKTANSQGRRKGRI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       : | : | | : | : | | : | | : | | : | | : | | : | | : | : | | : | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | 
                                                                                                                                                                                                                                                                                                                                                                                                 QASTPPQTPQP--NPPPVQATPHPPPA--VTPDLIVQTPVMTVVPPQPLQTPPVPPQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                  162 LVPPRLSKEELIQNMDRVDREITMVEQQISKLKKKQQQ------LEEEAAKPPEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              257 POPPAPAPOPVOSHPPI---LAATPOPVKTKKGVKRKADTTTPTTIDPIHEPPSLPPEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     E-----KPVSPP----PIESKHRSLVQIIYDENRKKAEAAHRILEGLGPQVELP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RVGQRGSGLSMSAARSEHEVSEIIDGLSEQENLEKQMRQLAVIPPMLYDADQQRIKFINM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -EKHKRKEEVEENKKSKAKAPPPKKTKKNNSSNSNVSKKEPAPMKSKPPP------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ----TYESEEDKCKPMSYEEKRQLSLDINKLP--------GEKLGRVVHI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           641 VGSKTVSQCKNFYFNYKKRONLDEI-LQQHKLKMEKERNARR-----KKKKAPAASEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         637 IQSREPS-----LKNSNPDEIEIDFETLKPSTLRELERYVTSCLRKKRKPQAEKVD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          694 AAFPPVVEDEEMEASGVSGNEEEMVEEAEALHASGNEVPRGECSGPATVNNSSDTESIPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PHTEAAKDTGQNGPK------PPATLGADGPPPGPPTPPRRTSRAPIEPTPASEA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 OOddddd-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TGAPTPPPAPPS-----PSAPPPV----VPKEEKEETAAAPPVEEGEEQKPPAAE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              773 QQQPPPPPPPPPSMPQQAAPAMKSSPPPIATQVPVLEPQLPGSVFDPI--GHFTQP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -----PTKVVAPPSSSDSSSDSDSDS-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          500 TDDSEEERA-----QRLAELQEQLKAVHEQLAALSQPQQNKP----KKKEKDKKEKKK-
                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                    Indels 511;
                                                                                                                                                                                                                                                Length 1362;
                                                                                                                                                                                       D52EFE1CF9960907 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     453 RF------PAVPDEPEEPVVAVSS-------PAVPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           723 РКЅКККСНРСКЕОККННИННООМООАРАРV-РООРРРООРР
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   314 KTTKLGQRRESSRPVKPPKKDVPDSQQHPA------
                                                                                                                                                                                                                                                                          ilarity 19.1%; Pred. No. 1.1e-05;
Conservative 171; Mismatches 502;
                                                                                                                                                                                                                                                   DB 1;
                                                                                                                                                                          (IN REF.
                                                                                                                                                                                                                                                   Score 382.5;
                            POLY-PRO.
POLY-PRO.
POLY-PRO.
POLY-GLN.
POLY-ALA.
POLY-ARG.
EM -> GP
                                                                                                                                                                                                                                                   2.9%;
                                                                                                                                                                                                  152219
                       986
1014
1033
1300
1308
1338
                                                                                                                                                                                               1362 AA;
                                                                                                                                                                                                                                                                          Similarity
                                                                        1028
1283
1301
1335
                                                                                                                                                                                                                                                                                                 Matches 280;
                                                                                                                                                                          CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         353
                                                                                                                                                                                                                                                                                                                                                                                                 201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          476
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          549
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            598
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                827
                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                               Local
  DOMAIN
DOMAIN
DOMAIN
                                                                                                  DOMAIN
                                                                           DOMAIN
                                                                                                                                                   DOMAIN
  SHIHHHH
                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              용
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             음. 성
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ઠ
                                                                                                                                                2385
                                                                                                                                                                                                  -----CPHGWISL 1577
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation -
1451 HSQTTAIPSCPEGTVPLYSGFSFLFVQGNQRAHGQDLGTLGSCLQRFTTMPFLFCN---- 1506
                                                                                             -------VNDVCNFASRNDYSYWLSTPALMPMNMAPITGRALEPYISRCTVCEGP 1554
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   the European Bioinformatics Institute. There are no restrictions on its most by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                              SNSAMVKSKKQEINKKLNTHNRNEPEYNISQPGTEIFNMPAITGTGLMTYRSQ----
                                                                                                                                                --AVQEHASTNMGLEAIIRKALMGKYDQWEESPPLSANAFNPLNASASLPAAMP---ITA
                                                                                                                                                                                                                                              2386 ADGRSDHTLTSPGGGGKAKVSGRPSSRKAKSPAPGLASGDRPPSVSSVHSEGDCN 2440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE=21590020; PubMed=11733348; French C.A., Miyoshi I., Aster J.C., Kubonishi I., Kroll T.G., Dal Cin P., Vargas S.O., Perez-Atayde A.R., Fletcher J.A.; "BRD4 bromodomain gene rearrangement in aggressive carcinoma with
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                               1578 WKGFSFIMFTSAGSEGTGQALASPGS-----CLEEFRASPFLECHGRGTCN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
-!- SUBCELLULAR LOCATION: Nuclear (Potential).
-!- SIMILARITY: Contains 2 bromodomains.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    16-OCT-2001 (Rel. 40, Created)
28-FEB-2003 (Rel. 41, Last sequence update)
15-WAR-2004 (Rel. 43, Last annotation update)
Bromodomain-containing protein 4 (HUNK1 protein).
                                                                                                                                                                                                                                                                                                                                                                                                                           PRT; 1362 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BROMODOMAIN 1.
BROMODOMAIN 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR001487; Bromodomain.
Pfam; PF00439; bromodomain; 2.
PRINTS; PR00503; BROMODOMAIN.
SMART; SW00297; BROMO; 2.
PROSITE; PS50014; BROMODOWAIN 1; 1.
PROSITE; PS50014; BROMODOMAIN 2; 2.
Bromodomain; Repeat; Nuclear protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POLY-SER.
POLY-HIS.
POLY-PRO.
POLY-PRO.
POLY-GLN.
POLY-PRO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LYS-RICH.
SER-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       translocation t(15;19)."; ... Am. J. Pathol. 159:1987-1992(2001)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; AF386649; AAL26987.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE OF 1-722 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; Y12059; CAA72780.1;
HSSP; Q92831; 1B91.
Genew; HGNC:13575; BRD4.
                                                                                                                                                                                                                                                                                                                                                                                                                           STANDARD;
                                                                                                                                                                                               1555 AIAIAVHSQTT---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    rissum-Placenta;
                                                                                                                                                                                                                                                                                                                                                                                                                        BRD4 HUMAN
O60885; Q96PD3;
16-OCT-2001 (Re)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BRD4 OR HUNK1
                                              2278
                                                                                                                                              2331
                                                                                               1507
                                                                                             셤
                                                                                                                                         ò
                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                          g
                                              ò
                                                                                                                                                                                                                                              ઠે
```

412

687

693

722 802 772 849

```
1097 RAASVVQPQPLVVVKEEKIHSPIIRSEPFSPSLRPBPP-----KHP------ESI 1140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1130 SVQLHVPYSEHAKAPVGPVTMGLPLPMDPKKLAPFSGV----KQEQLSPRGQAGPPESLG 1185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1186 VPTAQEASVLRGTALGSVPGGSITKGIPST----RVPSDSAITYRGSITHGTPADVLYK 1240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1196 IKN-------MGS--WASLVQKHPTTPSSTAKSSDSFEQFRRAAREKEEREKALK 1242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1241 GTITRIIGEDSPSRLD--RGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSS 1298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ------EDALEOA 1272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GPPHETAAPKRIYDMMEGRVGRAISSASIEGLMGRAIPPERHSPHHLKEOHHIRGSITQG 1358
                                                                                                                                                                                                                                                                                                                               982 РОРРРОООНОРРЯКРИНДРМОРЯТНІООРРРРОССООРРРРСООРРРРООРАКРООЙІО 1041
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1042 HHHSPRHHKSDPYS----TGHLREAPSPLMIHSPQMSQFQSLTHQSPPQQNVQPKKQEL 1096
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         --PTISNPPPLISSAKHPSVLERQIGAISQGM 1129
                                                                                                                                                                                                                                                                 PGSSPRGKSRSPAPPADKEAF-----AAEAQKLPGDPPCWTSGLPFPVPPREVIK 1064
                                             961
GAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTP----TGDPRANASP----QKPL
                                                                                                                                                                                                  931 ҮLQQLQKVQPPTPLLPSVKVQ------SQPPPPLPPPPPPPPPPPPPP
                                                                                                                                    DLKQLKQRAAAIPPIQVTKVHEPPREDAAPTKPAPPAPPPPP-----QPESDAPQQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
MEDLINE-96290553; PubMed=8703835;
Nilson I., Loechner K., Siegler G., Greil J., Beck J.D., Fey G.H.,
Marschalek R.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              "Exon/intron structure of the human ALL-1 (MLL) gene involved in translocations to chromosomal region 11q23 and acute leukaemias."; Br. J. Haematol. 93:966-972(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1273 RRAHEEARRRQEQQQQRQEQQQQQQQAAAVAAATPQAQSSQ------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Bukaryota, Metazoa, Chordata, Craniata, Vertebrata, Buteleostomi,
Mammalia, Butheria, Primates, Catarrhini, Hominidae, Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
MEDLINE=93046667; PubMed=1423624;
MEDLINE=91046667; Cleary M.L.;
"Involvement of a homolog of Drosophila trithorax by 11q23 chromosomal translocations in acute leukemias.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 protein)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      003164; 013743; 013744; 014845; 016364; 09UMA3; 01-0CT-1993 (Rel. 27, Created) 01-NOV-1995 (Rel. 32, Last sequence update) 29-FEB-2003 (Rel. 41, Last sequence update) Zinc finger protein HRX (ALL-1) (Trithorax-like) MLL OR HRX OR ALLI OR TEXI OR HTRX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3969 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1243 AQAEHAEKEKERLRQERMRSRED----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |:| : | | : | | 1317 -PQSMLDQQRELARKREQERRRE 1339
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1359 IPRSYVEAQEDYLR-REAKLLKRE 1381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE OF 1-1909 FROM N.A.
MEDLINE=93390935; PubMed=8378076;
                                                                     884 PARPPAVSPALTOTP-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Cell 71:691-700(1992)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1091 -TARPVLPRP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1299
                                                                                                                                                                                                                                                                           1015
                                                                                                                                                                                                                                                                                                                                                                                                               1065
                                                                                                                                           962
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC A PAC 
                                                                                                                                                                                                                                                                                                                        음
                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                         ò
                                                                                                                                                                                        a
                                                                                                                                                                                                                                                      ð
                                                           g
```

```
INTERACTION WITH SBF1.
MEDILINE-591967760, Pubmed-9537414;
MEDILINE-591967760, Pubmed-9537414;
Cui X., De Vivo I., Slany R., Miyamoto A., Firestein R., Cleary M.L.;
"Association of SET domain and myotubularin-related proteins modulates
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHROWOSCMAL THANSLACTANCE.

MEDLINE-20183971; PubMed=10706619;
Megonigal M.D., Cheung N.-K.V., Rappaport E.F., Nowell P.C.,
Milson R.B., Jones D.H., Addya K., Leonard D.G.B., Kushner B.H.,
Williams T.M., Lange B.J., Pelix C.A.,
Williams T.M., Lange B.J., Pelix C.A.,
"Detection of leukemia-associated MLL-GAS7 translocation early during
                                                                                                                                                                                                                                                                                          SEQUENCE OF 1251-1538 FROM N.A.
MEDIATNE=94215165; PubMed=8162375;
MEDIATNE=94215165; PubMed=8162375;
Gu Y., Alder H., Nakamura T., Schichman S.A., Prasad R., Canaani O.,
Saito H., Croce C.M., Canaani E.;
"Sequence analysis of the breakpoint cluster region in the ALL-1 gene
involved in acute leukemia.";
cancer Res. 54:2326-2330(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE OF 1212-1603 FROM N.A.
MEDLINE=95315013; PubMed=7794749;
Marschalek R., Greil J., Lochner K., Nilson I., Siegler G.,
Zweckbronner I., Beck J.D., Fey G.H.;
"Molecular analysis of the chromosomal breakpoint and fusion
transcripts in the acute lymphoblastic SEM cell line with chromosomal
translocation t (4;11).";
Yamamoto K., Seto M., Komatsu H., Iida S., Akao Y., Kojima S., Kodera Y., Nakazawa S., Ariyoshi Y., Takahashi T., Ueda R., "Two distinct portions of LTG19/ENL at 19p13 are involved in t(11,19)
                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE OF 1251-1654 FROW N.A. (ISOFORM 14P-18B).
MEDLINE=95322025; PubMed=7598802;
Mbangkollo D., Burnett R., McCabe N., Thirman M., Gill H., Yu H.,
Rowley J.D., Diaz M.O.;
"The human MLL gene: nucleotide sequence, homology to the Drosophila
trx zinc-finger domain, and alternative splicing.";
DNA Cell Biol. 14:475-483(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE OF 1421-1540 FROM N.A.
MEDLINE=94020842; PubMed=8414518;
Forster A., Rabbitts T.H.;
"A method for identifying genes within yeast artificial chromosomes: application to isolation of MLL fusion cDNAs from acute leukaemia translocations.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   mixed
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -:- FUNCTION: Possibly acts as a transcriptional regulatory factor.
-:- SUBGNIT: Interacts with SBF1.
-:- SUBCELIULAR LOCATION: Nuclear.
                                                                                                                                                                                      Evans G.A.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHROMOSOWAL TRANSLOCATION WITH AF3P21.
MEDLINE-2011s194; PubMed=10648423;
Sano K., Haydkawa A., Piao J.-H., Kosaka Y., Nakamura H.;
Navel SH3 protein encoded by the AF3p21 gene is fused to the inneage, leukemia protein in a therapy-related leukemia with
                                                                                                                                                                                    Djabali M., Selleri L., Parry P., Bower M., Young B.D., E. A trithorax-like gene is interrupted by chromosome 11q23 translocations in acute leukaemias.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          chemotherapy with DNA topoisomerase II inhibitor
Proc. Natl. Acad. Sci. U.S.A. 97:2814-2819(2000)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHROMOSOMAL TRANSLOCATION WITH GAS7.
                                                                                                                                                                   MEDLINE=93265134; PubMed=1303259;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Br. J. Haematol. 90:308-320(1995)
                                                                                                                             SEQUENCE OF 1317-2328 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          growth control.";
Nat. Genet. 18:331-337(1998).
                                                                                                                                                                                                                                                    Nat. Genet. 2:113-118(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Oncogene 8:3157-3160(1993).
                                                                                     Oncogene 8:2617-2625(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                t(3;11)(p21;q23).";
Blood 95:1066-1068(2000).
                                                                                                                                                       TISSUE=Brain;
                                                                        eukemia.
```

```
셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                                                                                                                  ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                         Name=14F-18b;
Isold=003164-2; Sequence=VSP_006666;
Isold=003164-2; Sequence=VSP_006666;
Isold=003164-2; Sequence=VSP_006666;
Isold=003164-2; Sequence=VSP_0offend and T and B lymphocytes.
-!-DISEASE: Involved in acute leukemias by chromosomal translocations t(11;19) (q23;p13.3) that involves MLL and MLLTI/AR4; t(9;11) (q22;q23) that involves MLL and MLLTI/AR4; t(9;11) (q22;q23) that involves MLL and MLLTI/AR4; t(9;11) (q13;q23) that involves MLL and MLLTI/AFX1; t(10;11) (q13;q23) that involves MLL and MLLTI/AFX1; t(10;11) (q13;q23) that involves MLL and MLLTI/AFX1; t(11;19) (q23;p23) that involves MLL and SLL; t(11;19) (q23;p23) that involves MLL and SLL; t(11;19) (q23;p23) that involves MLL and SLL; t(11;19) (q23;p23) that involves MLL and SLL; t(11;19) (q23;p23) that involves MLL and SLL; t(11;19) (q23;p23) that involves MLL and SLL; t(11;19) (q23;p23) that involves MLL and SLL; t(11;19) (q23;p23) that involves MLL and SLL; t(11;19) (q23;p23) that involves MLL and SLL; t(11;19) (q23;p23) that involves MLL and SLL; t(11;19) (q23;p23) that involves MLL and SLL; t(11;19) (q23;p23) that involves MLL and SLL; t(11;19) (q23;p23) that involves MLL and SLL; t(11;19) (q23;p23) that involves MLL and SLL; t(11;19) (q23;p23) that involves MLL and SLL; t(11;19) (q23;p23) that involves MLL and SLL; t(11;19) (q23;p23) that involves MLL and SLL; t(11;19) (q23;p23) that involves MLL and SLL; t(11;19) (q23;p23) that involves MLL and SLL; t(11;19) (q23;p23) that involves MLL and SLL; t(11;19) (q23;p23) that involves MLL and SLL; t(11;19) (q23;p23) that involves MLL and SLL; t(11;19) (q23;p23) that involves MLL and SLL; t(11;19) (q23;p23) that involves MLL and SLL; t(11;19) (q23;p23) that involves MLL and SLL; t(11;19) (q23;p23) that involves MLL and SLL; t(11;19) (q23;p23) that involves MLL and SLL; t(11;19) (q23;p23) that involves MLL and SLL; t(11;19) (q23;p23) that involves MLL and SLL; t(11;19) (q23;p23) that involves MLL and SLL; t(11;19) (q23;p23) that involves MLL and SLL; t(11;19) (q23;p23) that involves MLL and SLL; t(1
                                                                                                                                                                                                                                                                                                                                         This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
            Event=Alternative splicing; Named isoforms=2;
                                      IsoId=Q03164-1; Sequence=Displayed;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        JOINED.
JOINED.
JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              JOINED.
JOINED.
JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                    L04284; AAA58669.1; -. Z69744; CAA93625.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CAA93625.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CAA93625.1;
CAA93625.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CAA93625.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CAA93625.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CAA93625.1;
CAA93625.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CAA93625.1;
BAA03407.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CAA93625.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAA92511.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CAA93625.1;
CAA93625.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CAA93625.1;
CAA93625.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CAA93625.1;
CAA93625.1;
 ALTERNATIVE PRODUCTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CAA93625.1
CAA93625.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CAA93625.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CAA93625.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CAA93625.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CAA93625.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CAA93625.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CAA93625.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CAA93625.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CAA93625.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CAA93625.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CAA93625.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CAA93625.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CAA93625.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Z69748;
Z69749;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Z69764;
Z69765;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Z69751;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       269752;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   269754;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 269760;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          269761;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         269762;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           :991692
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          169767;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     269768;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1691691
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              569770;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         569773;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              L01986;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   69775;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               : 44692
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   D14540;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Z69747
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL;
EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL;
EMBL;
EMBL;
EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL;
EMBL;
EMBL;
EMBL;
EMBL;
EMBL;
EMBL;
EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL;
EMBL;
EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL;
 <del>-</del>
```

```
-----KOTDQPKAQGQESDSSETSVRGPRIKHVCRRAAVALGRKRAVFPDDMPTLSAL 1096
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 286 RRNHARKQWKQKFCQRYDQLMEALEKKVERIENNPRRRAKESKVREYYEKQFPEIRKQRE 345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           905 LYPVGRVSKEKVVGEDVATSSSAKKATGRKKSSSHDSGTDITSVTLGDTTAVKTKILIKK 964
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         441 REKFMOHPKNFGLIASFLERKTVAECVLYYYLTKKNENYKSLVRRSYRRGKSOQQQQQ 500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 226 LVQIIYDENRKKAEAAHRILEGLGPQVELPLYNQPSDTRQYHENIKINQAMRKKLILYFK 285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              862 - RDADKSVEKDK-----SRERDREREKENKRESRKEKR-----KKGSEIQSSSA 904
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ---TPQQSAELASMELNESSRWTEEEMETAKKGLLEHGRNWSAIARMVGSKTVSQCKNFY 653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----FPWFTPGSQTERGRNKDKAPEELSKD-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              L------OERMOSRVGORGSGLSMSAAR-SEHEVSEIIDG------LS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              .g. 540 --NDKEDLLKEKTDDTSGEDNDEKEAVASKGRKTANSQGRRKGRITRSMANEANSEAI-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   51 SHLSPGSIIQPQRRRPSLLS-EFQPGNERSQELHLRPESHSYLP----ELGKSEMEFIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               680 SPLHSGTRFDMHKRSPLLRAPRFTP----SBAHSRIFESVTLPSNRTSAGTSSSGVSNR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                106 KRPRLELLPDPLLRPSPLLATGQPAGSEDLTKDRSLTGKLEPVSPPSPPHTDPELELVPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             787 -LATSALNPTFTFPSHSLT----QSGESABKNQRPRKQTSAPAEPFSSSSPTPL----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            381 EQENLEKQMRQLAVIPPMLYDADQQRIKFINMNGLMADPMKVYKDRQVMNMWSEQEKETF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         501 QQQQQQQQPMPRSSQEEKDEKEK---------EKEAEKEEEKPEVE--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
2.9%; Score 380; DB 1; Length 3969;
Best Local Similarity 17.9%; Pred. No. 3.7e-05;
Matches 589; Conservative 357; Mismatches 1073; Indels 1270; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                166 RLSKEELIQNMDRVDREITMVEQQISKLKKKQQQLEEEAAKPPEPEKPVSPPIESKHRS
                                                                                                                                                                                                                                                          MIM; 159555, -
O GO:0005634; C:nucleus; TAS.
GO; GO:0005634; E:RNA polymerase II transcription factor a
GO; GO:0003702; F:RNA polymerase II transcription factor a
GO; GO:0003366; P:transcription from Pol II promoter; TAS.
InterPro; IPR001487; Bromodomain.
InterPro; IPR003889; FYrich. C.
InterPro; IPR003889; FYrich. N.
InterPro; IPR003616; PostSET.
                                                                                                                         EMBL; AF231998; AAG26332.2; ALT_TERM.
PIR; A44265, A44265.
TRANSFAC; T02337; -.
Genew; HGNC:7132; MLL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR002857; Znf CXXC.
InterPro; IPR001965; Znf PHD.
Pfam; PP00178; AT hook; Z.
Pfam; PP00628; PHD; 3.
Pfam; PP02008; SET; 1.
SMART; SM00297; BR0MO; 1.
SMART; SM00597; BR0MO; 1.
EMBL; S78570; AAB34770.1; -.
EMBL; X83604; CAA58584.1; -.
EMBL; S66422; AAB28545.1; -.
EMBL; AF231999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR001214; SET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        965
```

3053SPGPSQISNAAV	SPTESSSKESQUIAEIIS-PESPDRPPHSQTSGSCYYHVISKVPRIRTPSYSPTQRSPGC	2 A
2083 EGELRPKQPGPVKLGGEAA	SYVEAOEDYLREEAKLIKREGTPPPPPSRDLTFAYKTOALGPLKKPAH	8 8
3018 NNQDLTRNS	1302 HETAAPKRIYDMMEGRVGRAISSASIEGIMGRAIPPERHSPHHIKEQHHIRGSIIQGIPR 1361	ે દે
2024 EKTQSKPFSIQELELRŞLG	ISLRRKFLNGLEPENIHAMIĠSMTIDCLGILNDLSDCEDKLFPIGYQĊŚ	q
	KKGHVLSYEGGMSVT	δ
	1954YHFWCSRAKNCVFLDDKKVYCQRHRDLIKGEVVPENGFEVFRRVFVDFEG 2003	qq
1937 LIPERBARNARFERFRAD 1	1231 HGTPADVLYKGTITRIIGEDSPSRLDRGREDSLPKGHVIYEG 1272	ò
2863 LGESPESSSSELLNLGEGL		7 A
1907 VRPAATFPPATHCPLGG	TISDALITAGI BALTAKGI BALTAKGI BALTAV PSDSALITARGELT	ò
2808RRVHTSTPSDKNLL	1103 FESCHAZEZIONENGRASFESSI CONFERENCE IN STATEMENT STA	÷ 6
1883 MKGIITAVEPSKPT	AVERSONDINIAQMQEREBINGRIEQFFININILIFAFNERGGEGEFUSFIFFEFFFF	3 8
2749 GTENLKIDRPEDAGEKEHV	1112 AKHPSVLBRQIGA-ISQGMSVQLHVPYSEHAKAPVGPVTMGLPLPMDFKKLA 1162	ð í
1856	DGGQPEIKKANSMVKSFFIRQMERVFPWFSVKKSRFWEPNKVSSNSGMLPN	QQ
	1111	λō
1844 - GGGSSSRPASHS	1700 QDDQQPLDLEGVKRKMDQGNYTSVLERSDDIVKITQAAINS 1740	q
: FGLTPLYGVRSYG	1010 DAPQQPGSSPRGKSRSPAPPADKEAFAAEAQKLPGDPPCWTSGLPFPVPREVIKASPHA 1069	ò
	1699	qq
1/68FSSKASSSFUSTER 	964 KQLKQRAAAIPPIQVTKVHEPPREDAAPTKPAPPAPPPQNLQPES 1009	ò
COCO VIVIDIFICATIONS COLUMN	1585 MMQCGKCDRWVHSKCENLSDEMYEILSNLPESVAYTCVNCTERHPAEWRLALEKELQISL 1644	qq
NACO CONTRACTOR TO THE COST OF THE COST OST OF THE COST OF THE COST OF THE COST OF THE COST OF THE COST OST OF THE COST OF THE COST OF THE COST OF THE COST OF THE COST OST OF THE COST OF THE COST OF THE COST OF THE COST OF THE COST OS	963	ζ
LKPEFMUEV	1525 WICTKCVRCKSCGSTTPGKGWDAQWSHDFSLCHDCAKLFAKGNFCPLCDKCYDDDDYESK 1584	QΩ
	898GGSGRATTAKSSGAPQDSDSSATCSADEVDEAE 930	λ
	: : DD 1465 RPLEDQLENWCCRRCKFCHVCGRQHQATKQLLECNKCRNSYHPECLGPNYPTKFTKKKV 1524	qa
1676TAALENROTIINDYI : : : :::	873 GPAKGKDAEAABATAEGALKAEKKE 897	ζ
1	FVYCQVCCEPFHKFCLEE-NE 1464	qa
PRGIPLDAA	848ABELAVDIGKAEBPVKSECTEEAEE 872	'n
	1346 KVAPRPSIPVRQKPKEKEKPPPVNKQENAGTLNILSTLSNGNSSKQKIPADGVHRIRVDF 1405	qa
15/8 KASQDRKLISTPREIAKSP	847	ò
	1345	q
1531 PELGKPROSPLTYEDHGAPI 	TGONGPKPPATLGADGPPGPPTPPRRTSRAPIEPTPASEATGAPTPPAPPSPS	ò
2196 STSSLSPORSKLI	::: : : : : : : : : : : : : : :	; qq
1472 DVRSLIGSPGRTFPPVHPL	REMUREARING MILES TO THE STATE OF THE STATE	3 8
2160 RPLPSAG	CALLOL EXCENDENCE CONTRACTOR CONT	3 5
1412 EGLVATVKEAGRSIHEIPRI	VQ	ė

------HAHQHSPISPRTQDALQQRPSVLHNTG 1882 TIARTPAKNLAPHH------ASPDPPAPPASASDPHR 2023 ---PTPEGHMTPDHFIQGHMDADHISSP----PCGSVEQGHG 3017 GYHGSSYSPEGVE-PVSPVSSPSLTHDKGLPKHLEELDKSHL 2082 | |: ||| | : ||| 3052 VQTTPPHLKPATEKLIVVNQNMQPLYVLQTLPNGVTQKIQLT 3105 TSSNLQRTVVTVGNKNSHLDGSSSSEMKQSSASDLVSKSSSL 2303 | :: | | :: | | :: | | AHNVAYPGIP-KLAPQVH---NTTSRELNVSKIGSFAEPSSV 2357 ::|||: ----SFPHLH----LRGQRNDRDQHTDSTQSANSSPDEDTEV 2399 GIIDLSQVPHLPVLVPPT-PGTPATAMDRLAYLPTA-PQP-- 1767 ----PGGPTHLTKPTTTSSSERERDRDRERDRDREREKSIL 1815 FYSSSTGKKRGKRSAEGQVDGADDLSTSDEDDLYYYNFTRTV 2688 EQCDLPKISQLDGVDDGTESDTSVTATTRKSSQIPKRNGKEN 2748 : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | : | | : | : | | : | : | | : | : | | : | : | | : | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : LDTYNTELLKSDSDNNNSDDCGNILPSDIMDFVLKNTPSMQA 2862 LGLDSNREKDMGLFEVFSQQLPT-TEPVDSSVSSSISAEEQF 2921 TGHAFLAKPPARSGLEPASSPSK--------1977 ---SDLSVLTTRSPTVPSQNPSRLAVISDSGEKRVTITEKSV 2967 AHL--PHLRPLPES----QPSSSPL--LQTAP-GVKGHQRVV 2133 2134 TLAQHISEVITQDYTRHHPQ-----QLSAPLPAPLYSFPGASCPVL------ 2174 LDVMADARALERACYEE-SLKSRPGTASSSGGSIARGAPVIV 1530 PFA-----GHLPRGSPVTWREPTPRLQEGSLSSS 1577 PHSTVPEHHPHPISPYEHLLRGVSGVDLYRSHIPLAFDPTSI 1637 APNPTYPHLYPPYLIRGY-----1675 I-------1SQQMHHN------1 MGSSSRDRRQKGKKSCKETFKEKHSSKSFLEPGQVTTGEEGN 2459 -----VLRSTSTSP 1906 REELRHTPELPLAPRPLKEGSITOGTPLKYDTGASTTGSKKH 1471 IWRPGTEQSSGSSGG----

```
locus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CDNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hrm
                                                                                                                                                                                                                                                                                                                      8
2207
                                                                                                                             2295
                                                                                                                                                                          2355
                                                                                                                                                                                                 3394
                                                                                                                                                                                                                          2402
                                                                                                                                                                                                                                                                                             3453 QHVNQLLASKTGIHSSQRDLDSASG-----PQVSN------FTQTV--DAPN 3491
                                                                                                                                                                                                                                              OPVALPP-SSGMFPQLGTSQT-PSTAAİTAASSICVLPSTQTTGITAASPSGEADEHYQL 3452
SSVSSTPSVMETNTSVLGPMGGGLTLTTGLNPSLPTSQSLFPSASKGLLPMSHHQHLHSF
                                                3166 PAATQSSFPPNISNPPSGLLIGVQPPPDPQLLVSESSQRTDLSTTVATPSSGLKKRPISR
                                                                                                                                          ::: | |::: | |::: | 3286 IKRSKSSIMYFEPAPLLPQSVGGTAATAAGTSTISQDTSHLTSGSVSGLASSSSVINVVS
                                                                                                                                                                                                                          QWEESPPLSANAFNPLNASASLPAAMPITAADG-----RSDHTLTSPGGGGK----
                                -----SEGGKRSP----SEGGKRSP----SEGGKRSP---
                                                                                -----SRSAVYPLL
                                                                                                     LOTRKNKKLAPSSTPSNIAPSDVVSNMTLINFTPSQLPNHPSLLDLGSLNTSSHRTVPNI
                                                                                                                             2246 YRDGEQT----SPSRMGSKSPGNTSQPPAFFSKLTESNSAMVK-----SKKQEINKKLN
                                                                                                                                                                           2296 THNRNEPEYNISQPGTEIFNMPAITGTGLMTYRSQAVQEHASTNMGLEAIIRKALMGKYD
                                                                                                                                                                                                3346 MOTITIPISSASVPGHVILINPRILGIPDIGSISNLLIKASQQSLGIQ-----D
                                                                                                                                                                                                                                                                       ----AKVSGRPSSRKAKSPAPGLASGDRPPSVSSVHSEGDCNRRTPLTNRVWEDRPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Zhou J., Hertz J.M., Leinonen A., Tryggvason K.; Complete amino acid sequence of the human alpha 5 (IV) collagen chain and identification of a single-base mutation in exon 23 converting glycine 521 in the collagenous domain to cysteine in an
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Homo sapiens (Human).
Bukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia, Butheria, Primates, Catarrhini, Hominidae, Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TISSUE=Placenta;
MEDLINE=90337990; PubMed=2380186;
MEDLINE=90337990; PubMed=2380186;
Pihlajaniemi T., Pohjolainen E.R., Myers J.C.;
"Complete primary structure of the triple-helical region and the carboxyl-terminal domain of a new type IV collagen chain, alpha
                                                                                                                                                                                                                                                                                                                                    SMG---LEQN---KALSSAVQASPTSPG-GSPSSPSSGQRSASPSVPGP 3533
                                                                                                                                                                                                                                                                                                                     SAGSTPFPYNPLIMRLQAGVMASPPPPGLPAGSGPLAGPHHAWDEEPKP 2504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE=94165049; PubMed=8120014;
Zhou J., Leinonen A., Trygyvason K.;
"Structure of the human type IV collagen COL4A5 gene.";
J. Biol. Chem. 269:6608-6614(1994).
                                                                               -----EPGH---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE OF 1-910 FROM N.A., AND VARIANT AS CYS-521.
                                                                                                                                                                                                                                                                                                                                                                                                   CA54 HUMAN STANDARD; PRT; 1685 AA. P29400; Q16006; Q16126; Q1-DEC-1992 (Rel. 24, Created) PFB-1994 (Rel. 28, Last sequence update) 10-OCT-2003 (Rel. 42, Last annotation update) Collagen alpha 5(IV) chain precursor.
                                                                               ---EPNKTSVLGGGEDGIEPVSPPEGMT-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Alport syndrome patient.";
J. Biol. Chem. 267:12475-12481(1992)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Biol. Chem. 265:13758-13766(1990)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TISSUE=Kidney;
MEDLINE=92316923; PubMed=1352287;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE OF 924-1685 FROM N.A.
MEDLINE=91169491; PubMed=2004755;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE OF 85-1685 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=9606;
                                                                               2208
                                  2175
                                                                                                     3226
                                                                                                                                                                                                                          2356
                                                                                                                                                                                                                                                3395
                                                                                                                                                                                                                                                                       2403
                                                                                                                                                                                                                                                                                                                     2456
            3106
                                                                                                                                                                                                                                                                                                                                            3492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COL4A5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5 (IV).
                                                                                                                                                                                                                                                                                                                                                                                RESULT 40
                                                                                                                                                                                                                                                                                                                                                                                         음
                                                     g
                                                                            ò
                                                                                                  셤
                                                                                                                            ò
                                                                                                                                               요
                                                                                                                                                                         ð
                                                                                                                                                                                                원
                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                             유
                                  ઠે
                                                                                                                                                                                                                                                                                                                     ò
```

```
gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE=9203559; PubMed=1376965;
Knebelmann B., Deschenes G., Gros F., Hors M.-C., Gruenfeld J.-P.,
Tryggvason K., Gubler M.-C., Antignac C.;
"Substitution of arginine for alycine 325 in the collagen alpha 5
(IV) chain associated with X-linked Alport syndrome: characterization of the mutation by direct sequencing of PCR-amplified lymphoblast
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Cassiman J.-J.,
                                IV collagen alpha
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Tryggvason K.; "Single base mutation in alpha 5(IV) collagen chain gene converting conserved cysteine to serine in Alport syndrome.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE OF 1442-1471 FROM N.A.
MEDLINE=90252791; PubMed=2139699;
Myers J.C., Jones T.A., Pohjolainen E.R., Kadri A.S., Goddard A.D.,
Solomon E., Pihlajaniemi T.;
"Molecular cloning of alpha 5(IV) collagen and assignment of the ge
to the region of the X chromosome containing the Alport syndrome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Zhou J.,
Rascher W.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Cassiman J.-J., Marynen P.;
"Differential splicing of COLAAS mRNA in kidney and white blood cells: a complex mutation in the COLAAS gene of an Alport patient
                                                                                                                                                                                                                                                                                                                chain with
gene to the
                                                                                                                                                                   SEQUENCE OF 914-1685 FROM N.A.
MEDLINE=90160375; PubMed=1689491;
Hostikka S.L., Eddy R.L., Byers M.G., Hoeyhtyae M., Shows T.B.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VARIANT AS SER-1564.
MEDLINE-91169492; PubMed=1672282;
Zhou J., Barker D.F., Hostikka S.L., Gregory M.C., Atkin C.L.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŗ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE OF 1258-1270 FROM N.A. (ISOFORM 2).
MEDLINE-94133540; PubMed-8301933;
Guo C., van Damme B., van Damme-Lombaerts R., van den Berghe
                                                                                                                                                                                                                                                                                                                                                     gene to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REVIEW ON VARIANTS.
MEDLINE-97338662; PubMed-9195222;
Lemmink H.H., Schroeder C.H., Monnens L.A.H., Smeets H.J.M.;
"The clinical spectrum of type IV collagen mutations.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          converting glycine 325
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Renieri A., Seri M., Myers J.C., Pihlajaniemi T., Massella Rizzoni G.F., de Marchi M.;
De novo mutation in the COL4A5 gene converting glycine 325 glutamic acid in Alport syndrome.";
Hum. Mol. Genet. 1:127-129(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VARIANTS AS THR-1517; SER-1538 AND GLN-1563.
MEDLINE=94010948; PubMed=8406498;
Lemmink H.L., Schroder C.H., Brunner H.G., Nolen M.R.,
Tryggvason K., Haggsma-Schouten W.A.G., Roodvoets A.P.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Submitted (SEP-1994) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Vanrenterghem Y., Devriendt K.,
                                                                                                                                                                                                                                                                         Tryggvason K.;
"Identification of a distinct type IV collagen alpha restricted kidney distribution and assignment of its locus of X chromosome-linked Alport syndrome.";
Proc. Natl. Acad. Sci. U.S.A. 87:1606-1610(1990).
Zhou J., Hostikka S.L., Chow L.T., Tryggvason K.;
'Characterization of the 3' half of the human type
s gene that is affected in the Alport syndrome.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Am. J. Hum. Genet. 46:1024-1033(1990)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              J. Hum. Genet. 51:135-142(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VARIANT AS GLU-325.
MEDLINE=93244772; Pubmed=1363780;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      deletes the NC1 domain.";
Kidney Int. 44:1316-1321(1993)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Mutat. 9:477-499(1997)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE OF 1-20 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Guo C., van Damme B., '
Marynen P.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Genomics 9:10-18(1991)
                                                                                                     Genomics 9:1-9(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VARIANT AS ARG-325.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          fragments.
```

```
Similarity
                                                                                                                                                                                                                                                                                                                                            VARIANT AS ARG-822
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1013 QQPG-
                                                                                                                                                                                                                                                                            sequencing."
                                                                                                                                                                                                                                                                                                                                                                                                                                              380;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                912
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                943
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1087
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1147
                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VARIANTS AS VAL-420; 456-PRO--PRO-458 DEL; ASP-573; ASP-624; ASP-635; 802-GLT-PRO-4807 DEL, ARG-869; CYS-941; SER-1030; SER-1066; ASP-1143; ARG-1196; GLU-1261; SER-137 AND ARG-1649.

MEDLINE=99063529; PubMed=9848783; MAC-1649.

MARTIN P., Heiskari N., Zhou J., Leinonen A., Tumelius T., Hertz J.M., Barker D.F., Gregory M.C., Atkin C.L., Styrkarsdottir U., Neumann H., Springate J., Shows T.B., Pettersson E., Tryggvason K.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE=98112435; PubMed=9452056;
Neri T.M., Zanelli P., de Palma G., Savi M., Rossetti S., Turco A.E.,
Pignatti G.F., Galli L., Bruttini M., Renieri A., Mingarelli R.,
Trivelli A., Pinciaroli A.R., Raqaolo M., Rizzoni G.F., de Marchi M.;
"Missense mutations in the COL4A5 gene in patients with X-linked
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Knebelmann B., Breillat C., Forestier L., Arrondel C., Jacassier D., Giatras I., Drouot L., Deschenes G., Gruenfeld J.-P., Broyer M., Gublar M.-C., Antignac C.; Bepetrum of mutations in the COL4A5 collagen gene in X-linked Alport syndrome.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VARIANTS AS.

MEDIATE-96213754; PubMed=8651296;

MEDIATE-96.13754; PubMed=8651296;

Renieri A., Bruttini M., Galli L., Zanelli P., Neri T.M., Rossetti S.

Turco A.E., Heiskari N., Zhou J., Gusmano R., Massella L., Banfi G.,

Scolari F., Seesa A., Rizzoni G.F., Tryggvason K., Pignatti P.F.,

Savi M., Ballabio A., de Marchi M.,

"X-linked Alport syndrome: an SSCP-based mutation survey over all 51
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VARIANTS AS ARG-174; ARG-177; ARG-325; CYS-1410; TRP-1421; THR-1517
                                                                                                                    VARIANTS AS GLU-400; VAL-406; VAL-638; ALA-638; ARG-653; ARG-796; ARG-869; ARG-869; ARG-869; ARG-869; ARG-869; ARG-869; ARG-869; ARG-869; ARG-869; ARG-872 AND CYS-1241.

MEDLINE=55322976; PubMed=7599611;
Boye E., Flinter F., Zhou J., Tryggvason K., Bobrow M., Harris A.;
"Detection of 12 novel mutations in the collagenous domain of the COL4A5 gene in Alport syndrome patients.";
Hum., Mutat. S:197-204(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Barker D.F., Denison J.C., Atkin C.L., Gregory M.C.; "Common ancestry of three Ashkenazi-American families with Alport syndrome and COLAAS R1677Q.";
                                                                                                                 VAL-406; VAL-638; ALA-638; ARG-653; ARG-796; CYS-1241.
                                                                                                                                                                                                                                                                                                                                                                                                                          is
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         collagen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VARIANTS AS, AND VARIANTS ASP-430; SER-444; SER-619; ASN-664 AND
                                                                                                                                                                                                                                                                                                                                       MEDIANCE AS ANOTATION TO THE MEDIANCE AS A STATE OF THE MEDIANCE AS A STATE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE AND THE MEDIANCE AND THE MEDIANCE AND THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE OF THE MEDIANCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE=9623392; PubMed=8829632;
Tverskaya S., Bobrynina V., Tsalykova F., Ignatova M.,
Krasnopolskaya X., Evgrafov O.;
Kusnopolskaya X., Evgrafov O.;
Substitution of A1498D in noncollagen domain of a5(IV) collag
chain associated with adult-onset X-linked Alport syndrome.";
                         φ
                            gene
                    mutations in the COL4A5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 exons of the COL4A5 gene.";
Am. J. Hum. Genet. 58:1192-1204(1996)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        J. Hum. Genet. 59:1221-1232(1996)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Alport syndrome.";
Hum. Mutat. Suppl. 1:S106-S109(1998)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE=97094179; PubMed=8940267;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE=97295089; PubMed=9150741;
                                              patients with Alport syndrome.";
Genomics 17:485-489(1993).
                       'Identification of four novel
Smeets H.J.M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hum. Mutat. 7:149-150(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Genet. 99:681-684 (1997)
                                                                                                                                                                                                                                                                                                                             VARIANT AS ARG-1649.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VARIANT AS ASP-1498.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VARIANT AS GLN-1677.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ASP-1596
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AND
```

```
1189 AQEASVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYKGTITRIIG 1248
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               974 P-PIQVTKVHEPPRED-AAPTKPAPPAPPPQNLQPE------SDAP 1012
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1064 K-----ASPHAP-----1086
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ------PVTMGLP-----LPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGVPT 1188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1369 DYLRREAKLIKREGTPPPPPSRDLTEAYKTQALGPLKL-----KPAHEGLVATVKEA 1421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ----SSPRGKSRSPAPPADK-EAFAAEAOKLPGDPPCWTSGLP-FP---VPPREVI 1063
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    969
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   813
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         374 GFPGIQGPPGLPGPPGAAVMGPPGPPGFPGERGQKGDEGPPGISIPGPPGLDGQPGAPGL 433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     851
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          911
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -----SLLTPTGDP-----RANASPQKPLDLK-----QLKQRAAAI 973
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        494 GÓPGLPGLPGPPGSLGFPGQKGEKGQAGATGPKGLPGIP--GAPGAPGFPGSKGEPGDIL 551
"High mutation detection rate in the COL4A5 collagen gene in suspected Alport syndrome using PCR and direct DNA sequencing."; J. Am. Soc. Nephrol. 9:2291-2301(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1249 EDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETAAPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              QGPPGPPGIPGMKGEPGSIIMSSLPGPKGNPGYPGPPGIQGLPGPTGIPGPIGPPGPPGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         552 TFPGMKGDKGELGSPGAPGLPGLPGTPGQDGLPGLPGPKGEPGGTTFKGERGPPGNPGLP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GLHDTARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEHAKAPVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   612 GLPGNIGPM--GPPGFGPPGP-----VGEKGIQGVA-GNPGOPGIP-----GPKG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ----GLPGQPGLPGIPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SKGEPGI PĠIGLPGPPĠ----PKĠFPGIPGPPGAP------ĠTPGRIGLEĠ-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1309 RTYDMMEGRVGRAISSASIEGLMGRAIPPERHSPHHLKEQHHIRGSITQGIPRSYVEAQE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            203 MGPPGPPGLPGPKGNMGLNPQGPKGEKGEQGLQGPPGPPGQISEQKRPIDVEPQKGDQGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AVDICKAEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSSGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     263 PGDRGPPGPP----GIRGPPGPPGFKGEKGEQGEPGKRGKPGKDGENGQPGI---PGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  814 -SPSAPPVV------PKEEKEETAAAPPVEEGE--EQKPPAAEE----L
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ONGPKPPATLGADGPP-----PGPPTPPRRTSRAPIEPTPASEATGAPTPPPAPP--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                     VARIANTS AS GLU-579, LYS-633, ASP-947; VAL-953; ARG-1107; ARG-1158; SER-1170 AND TRP-1158.

MEDI-1170 AND TRP-1678, AND VARIANTS SER-444 AND ALA-739.

MEDI-118E=20030197; PubMed=10541141;

Inoue Y., Nishio H., Shirakawa T., Nakanishi K., Nakamura H.,

Sumino K., Nishiyama K., Iljina K., Yoshikawa N.;

"Detection of mutations in the CO1445 gene in over 90% of male patients with X-linked Alport's syndrome by RT-PCR and direct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PQDSDSSATCSADEVDEAEGGDKN----RLLSPRP----------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            680;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2.9%; Score 379; DB 1; Length 1685; 20.7%; Pred. No. 1.8e-05;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative 136; Mismatches 638; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    654 DPGQTITQPGKPGLPGNPGRDGDVGLPGDP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----GFKG-ALGPKGDRGFPGPPGPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Am. J. Kidney Dis. 34:854-862(1999).
```

```
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
             1525
                                                                                                                                                                                                                                                                                                                                                     1156
                                                                                                                                                                                                                                                                                                                                                                                                                                       1257
                                                                                                                                                                                                                                                                                                                                                                                                                                                            2028
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2085
                                                                                                                                                                                                                                                         1856
                                                                                                                1526 APVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSPVTMREPTPRLQEGSLSSSKASQDRKL 1585
                                                                                                                                                                                                                                                                                       -----VPPTPGTP-ATAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTHLTKPTTTSSSER 1796
                                                                                                                                                                                                                                                                                                                                                                          1857 AHQHSPISPRTQDALQQRPSVLHNTGMKGIITAVEPSKPTVLRSTSTSSPVRPAATFPPA 1916
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            G----QPGTRGLDGPPGPDGLQGPPGTSSVAHGF----LITRHSQTTD----APQ 1475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2261 SPGNTSQPPAFFSKLTESNSAMVKSKKQ-----------EINKKLNT 2296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1476 CPQGTLQVYEGFSLL-----YVQGNKRAHGQDLGTAGSCLRRFSTMPFMFCNINNVCNF 1529
           836
GRSIHEIPREE-------LRHTPELPLAPRPL-KEGSITQGTPLKYDTGAST
                                             1466 TGSKKHDVRSLIGSPGRTFPPVHPLDVMADARALERACYEESLKSRPGTASSSGGSIARG
                                                                                     ---GLP-----GPPGPM----DPNLLGSKGEKG---
                                                                                                                                                          1586 TSTPREIAKSPHSTVPEHHPHPISPYEHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLDA
                                                                                                                                                                              -----GQP-GL
                                                                                                                                                                                                  1646 AAAYYLPRHLAP--NPTYPH---LYPPYL------IRGYPDTAALENRQTI
                                                                                                                                                                                                                        SGQPGLPGPPGPKGNPGLPGQPGLIGPPGLKGTIGDMGFPGPQGVEGPPGPSGVPGQP--
                                                                                                                                                                                                                                            1686 INDYITSQQMHHNTATAMAQRADMLRGLSPRESSLALNYAAGPRGIIDLSQVPHLPVL--
                                                                                                                                                                                                                                                                                                                                 ERDRDRERDRDREREKSILTSTTTVEHAPIWRPGTEQSSGSSGSSGGSSSRPASHSH
                                                                                                                                                                                                                                                                                                                                                     ------
                                                                                                                                                                                                                                                                                                                                                                                              PGOPGPPGEKGKPGODGIPG---PAGOKG----EPGOP-----GPGNPGPPG--LPGL
                                                                                                                                                                                                                                                                                                                                                                                                                   1917 THCPLGGTLDGV--YPTLMEPVLLP--KEAPRVARPERPRADTGHAFLAKPPARSGLEPA
                                                                                                                                                                                                                                                                                                                                                                                                                                      SGQKGDGGLPGIPGNPGLPGPKGEPGFHGFPGVQGPPGPPGSPGPA-LEGPKGNPG--PQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                            SSPSKGSEPRPLVPP-VSGHATIARTPAKNLAPHHASPDPPAPPASASD---PHREKTQS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1258 GPPGRPGLPGPEGPPGLPGNGGI-----KGEKGNPGQPGLPGLPGLKGDQGPPGLQGNPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      KPFSIQELELRSLGYHGSSYSP--EGVEPVSPVSSPSLTHDKGLPKHLEELDKSHLEGE-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RP-----GLNGMKGDPGLPGVPGFPGMKGPS----GVP-----GSAGPEGEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2086 -- LRPKQPGPVKLGGEAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQRVVTLAQHISEVI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2144 TQDYTRHHPQQLSAPLPAPL--YSFPGASCPVLDLRRPPSDLYLPPPDHGAPARGSPHSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GGKRSPEPNKTSVLG-GGEDGIEPVSPPEGMTEPGHSRSAVYPLLYRDGEQTEPSRMGSK
                                                                                                                                                                                                                                                                                                           1099 SVGDPGLPGLPGTPGAKGQPGLPGFPGTP------GPPGPKGISGP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1351 GLIGP--PGPPGLPG----PSGQSIIIKGDAGPPGIPGQPGLKG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2297 HNRNEPEYNISQPGTEIFNMPAITGTGLMTYRSQ 2330
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1530 ASRNDYSYWLSTPEPMPMSMOPLKGOSIOPFISR 1563
                                                                                                                                      945 EP-
                                                 837
                                                                                                                                                                                                                        966
                                                                                                                                                                                                                                                                                     1744
                              1422
                                                                                                                                                                              896
                                                                                                                                                                                                                                                                  1054
                                                                                                                                                                                                                                                                                                                                 1797
                                                                                                                                                                                                                                                                                                                                                     1139
                                                                                                                                                                                                                                                                                                                                                                                              1157
                                                                                                                                                                                                                                                                                                                                                                                                                                                            1973
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2029
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1313
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1430
           787
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2202
                                                                 ઠે
                                                                                         g
                                                                                                                8
                                                                                                                                 원
                                                                                                                                                       ઠે
                                                                                                                                                                            g
                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                           Š
                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                  Š
                                                                                                                                                                                                                                                                                                                                                                                                                                 ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                      ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            음
                              ઠે
                                                                                                                                                                                                 ઠે
                                                                                                                                                                                                                                                                                     8
                                                                                                                                                                                                                                                                                                                                ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ઠે
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the BMBL outstation the European Bioinformatics Institute. There are no restrictions on its was by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to licenseeisb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            821
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TISSUE-Brain, Cerebellum, Hippocampus, and Substantia nigra;
MEDLINE=96081227; PubMed=8841849;
Schmitt I., Epplen J.T., Riess O.;
Predominant neuronal expression of the gene responsible for dentatorubral-pallidoluysian atrophy (DRPLA) in rat.";
Hum. MOI. Genet. 4:1619-1624(1995).
-!- SUBUNIT: Interacts with WWP1 and WWP2 (By similarity).
-!- TISSUE SPECIFICITY: Predominant neuronal expression, although markedly reduced amounts are found in most other tissues.
-!- DEVELOPMENTAL STAGE: Similar expression at all development stages (14.5 dpc, 17.5 dpc, newborns and adults).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TGONGPKPPATLGADGPPPGPPTPPRRTSRAPIEPTPASEATGAPTPPPAPPSPSAPPPV
                                                                                                                                                                                    Bukaryota, Metazoa, Chordata, Craniata, Vertebrata, Buteleostomi,
Mammalia, Eutheria, Rodentia, Sciurognathi, Muridae, Murinae, Rattus.
                                                                                                                                                                                                                                                                                                                                                                                                           gene)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 378; DB 1; Length 1183;
Pred. No. 1.4e-05;
5; Mismatches 604; Indels 486;
                                                                                                                                                                                                                                                                                                       TISSUE=Cerebellum, and Striatum;
MEDLINE=97317138; PubMed=91773996;
Loev S.J., Margolis R.L., Young W.S., Li S.-H., Schilling G.,
ASHWORTH R.G., Ross C.A.;
"Cloning and expression of the rat atrophin-I (DRPLA disease)
                                                                   01-OCT-1996 (Rel. 34, Last sequence update)
15-MAR-2004 (Rel. 43, Last annotation update)
Atrophin-1 (Dentatorubral-pallidoluysian atrophy protein)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TRONKDSMSMRSGRKKEAPGPREELRSRGRASPGGVSTSSSDG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             7FB9928DCADF9B1F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  T -> M (IN REF. 2
A -> V (IN REF. 2
MISSING (IN REF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REF.
Æ
1183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POLY-PRO.,
POLY-SER.
POLY-RRO.
POLY-HIS.
POLY-GIA.
POLY-ERO.
POLY-ERO.
POLY-PRO.
POLY-PRO.
POLY-PRO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ^
                                                                                                                                                                                                                                                                                                                                                                                                                             homologue.";
Neurobiol. Dis. 2:129-138(1995)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative 146;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; X89453; CAA61623.1; -.
InterPro; IPR002951; Atrophin.
Pfam; PF03154; Atrophin-1; 2.
                                             (Rel. 34, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; U31777; AAA80337.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             124778
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      20.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PR01222; ATROPHIN
STANDARD;
                                                                                                                                                                Rattus norvegicus (Rat)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       965 96
1183 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity
                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                              NCBI_TaxID=10116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  562
702
455
594
689
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              387
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     481
                                             01-OCT-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             321;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            m
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ઠે
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib.ch).
                                                                                                                                                                                                                                                                                                                                                                  1031 PLARLQMLNVT-----PHHQH----SHIHSHLHLHQQDAIHAASASVH-PLIDPL----- 1076
                                                                                                                                                                                                                                                                                                                                                                                                                      2115 SSSPLLQTAPGVKGHORVVTLAQ-----HISEVITQDYTRHH------PQQLSAP 2158
                                                                                                                                                                                                                                                                                                                                                                                                                                                             1077 ASGSHLTRIPYPAG-----TLPNPLLPHPLHENEVL-----RHOLFAAPYRDLPASLSAP 1126
                                                                              1888 TAVEPSKPTVLRSTSTSSPVRPAATFPPATHCPLG----GTLDGVYPTLMEPVLLPKEAP 1943
                                                                                                                                                                 1944 RVARPERPRADTGHAFLAKPPARSGLEPASS-PSKGSEPRPLVPPVSGHATIARTPAKNL 2002
                                                                                                                                                                                                                                                                                              982 FPFHPSLGPLERERLALAAGPALRPD------MSYAE-RLAAERQHAERVAALGND 1030
                                                                                                                                                                                                                                                                                                                                     PVSPVSSPSLTHDKGLPKHLEELDKSHLEGELRPKQPGPVKLGGEAAHLPHLRPLPESQP 2114
1828 RPGTEQSSGSSGSSGGGGSSSRPASHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGII 1887
                                                                                                                                                                                                                                                    2003 APHHASPDF------PAPPASASDPHREKTQSKPFSIQELELRSLGYHGSSYSPEGVE 2054
                                                                                                           TLSEYARPHVMSPGNRNHPF------YVPLGAVDPGLLGYNVPALYSSDPAARERB 927
                                                                                                                                                                                              2159 LPAPLYSFPGASCPVLDLRRPPSD----LYLPPPDHGA--PARGSPHSEGGKRSPEP 2209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      INTERACTION WITH FNBED.

INTERACTION WITH FNBED.

INTERACTION WITH FNBED.

MEDLINE-9713-117; PubMed-9171351;

Bedford M.T., Chan D.C., Leder P.;

"FBP WW domains and the Abl SH3 domain bind to a specific class of proline-rich ligands.";

"FBP WW domains and the Abl SH3 domain bind to a specific class of EMBO J. 16.2376-23831199;

-I FUNCTION: Transcriptional activator that binds to the AT-rich core sequence of the enhancer element of the AFP gene.

-I SUBGINIT: Interaction with FNBED.

-I SUBCELLULAR LOCATION: Woclear.

-I SIMILARITY: Contains 4 homeobox domains.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            [1]
SEQUENCE FROM N.A.
STRAIN-BALB/MX X ICR, TISSUE-Brain;
STRAIN-BALB/MX Y ICR, TISSUE-Brain;
STRAIN-BALB/MX N. ICR, TISSUE-Brain;
MEDLINE-96194902; PubMed-8654949;
Ido A., Miura Y., Watanabe M., Sakai M., Inoue Y., Miki T.,
Hashinoto T., Morinaga T., Nishi S., Tamaoki T.;
"Cloning of the CDNA encoding the mouse ATBFI transcription factor.";
Gene 168:227-231(1996).
                                         ---SLGPVPH----RPPFEPGS--AVATVPPYL-GPDTPALR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 [6-0CT-2001 (Rel. 40, Created)
16-0CT-2001 (Rel. 40, Last sequence update)
10-0CT-2003 (Rel. 42, Last annotation update)
Alpha-fecoprotein enhancer binding protein (AT motif-binding factor)
(AT-binding transcription factor 1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
VCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT; 3726 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; D26046; BAA05046.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
                                                844 CP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ABF1 MOUSE
                                                                                                                                                                                                                                                                                                                                            2055
                                                                                                                                878
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 42
ABF1_MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ATBF1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                          .연
                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   유
                                           셤
                                                                                       à
                                                                                                                                g
                                                                                                                                                                          ò
                                                                                                                                                                                                                                              - 8
```

```
2302
                    | | ||: :| | | 2127 PVEPLPADLAQLYQHQLNPTLLQQQNKRPRTRITDDQLRVLRQYFDINNSPSEEQIKEMA 2186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TKVHEPPREDAAPTKPAPP--APPPPQNLQPESDAPQQPGSSPRGKSRSPAPPADKEAFA 1036
                                                                                                        1630 HQTKARAAKLEAASGNSNGTGNSGGVSLSSSTPSPVGSSGANNT-----FTAT 1677
                                                                                                                                                                                1678 NPS------SAAMAPSVNALSQVPPESVVMPPLGNPISANIASPSEPKEAN 1722
                                                                                                                                                                                                                                                     ------VQAHLQQELQQQA 1769
                                                                                                                                                                                                                                                                                                                                                                                                  FYIPSAEFQLNPEVSLPVTSGALTLTGSGPGLLEDLKVQVQIPQQSHQQILQQQQQQQQL 1864
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ------EGAEGNTGPQE 1913
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -----ELVIQYN---ENKQK 1972
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -------PEPPP-----PPPPPPPPLPT--APPQPA 2066
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2187 D----KSGLPQKVIKHWFRNTLFKERQRNKDSPYNFSNPPITSLEELKIDSRPPSPEPQK 2242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2479
                                                                                                                                                                                                                                                                                                                          559 DEKEAVASKGRKTANSQGRRKGRITRSMANEANSEEAITPQQSAELASMELNESSRWTEE 618
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  863
                                                                         -----RKKAEAAHRILEGLGPQVELPLYNQ-PSDTRQYHENIKINQAMRKKLILYFKRR 287
                                                                                                                                                                                                                                                                                         AVIPPMLYDADQQRIKFINMNGLMADPMKVYKDRQVMNMWSEQEKETFREKFMQHPKNFG 452
                                                                                                                                                                                                                                                                                                                                                              453 LI---ASFLERKTVAECVLYYYLIKKN-----ENYKSLVRRSYRRGKSQQQQQQQQQQ
                                                                                                                                                                                                                                                                                                                                                                                                                                     504 QQQQQQPMPRSSQ--EEKDE---KEKEKEEKFEFKPEVENDKEDLLKEKTDDTSGEDN 558
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  619 EMETAKKGLLEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDEILQQHKLKMEKERN 678
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        871
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    --GSGRATTAKSSGAPQDSDSS---ATCSA 923
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         679 ARRKKKKAPAAASEEAAFPPVVEDEEMEASGVSGNEEEMVEEAEALHASGNEVPRGECSG
                                                                                                                                              NHARKQWKQKFCQRYDQLMEALEKKVER1ENNPRRR-------AKESKVREYY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ------DLAPGGGSEGTMLPPRIA------SD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2243 QEYWGSKRSSRTRFIDYQLRVLQDFFDANAYPKDDEFEQLSNLLNLPTRVIVVWFQNARQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DELATFNSKAEASDEKPKQADPPSAQPNQTQEKQGQPKPEMQQQLEQLEQKTNAPQP---
                                                                                                                                                                                                                   334 EKQFPEIRKQRELQERMOSRVGQRGSGLSMSAARSEHEVSEIIDGLSEQENLEKQM-RQL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -VVdd----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----PKEEKEETA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -----PP--AAEELAVDT-GKAEEPVK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ---KAE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               KARKNYENQGEGKDGERRELTNDRYIRTSNLNYQCKKCSLVFQRIFDLIKHQKKLCYKDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2363 DEEGODDSONEDSMDAMEILTPTSSSCSTPMPSOAYSTPAPSAAANTAPSAFLOLTAET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DEV----DEAEGGD-KNRLLSPRPSLLTPTGDPRANASPQKPLDLKQLKQRAAAIPPIQV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PATVNNSSDTESIP-SPHTEAAKDTGQNGPKPPATLGADGPPPGPPTPPRRTSRAPIEP-
 E-QQISKLKKKQQQLEEEAAKPPEPEKPVSPPPIESKHRSLVQIIYDEN-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ----TEEAE----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----EGPAKGKDAEAAEATAEGAL------
                                                                                                                                                                                                                                                                                                                                                                                                                                                          1723 RKKLADMIASRQQQQQQQQQQQQQTLAQAQAQ----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -TPASEATGAP--TPPPAPPSPSAP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAPPVEEGEEQK------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2025 FAKOYREHYDKLYPLRPQT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          864 SEC-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SLPDASKAKEKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1946 ARGNATKALLE-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      895 KKEG
                                                                                                                                              288
                                                                                                                                                                                                                                                                                         393
                                                                                                                                                                                                                                                                                                                                                                                                    1805
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1914
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                739
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2303
                                                                         235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      797
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             823
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    833
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 872
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            924
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   979
                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    쉽;
                               엄
                                                                                                   g
                                                                                                                                                                              ద
                                                                                                                                                                                                                                                                                                                          유
                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          엽
                                                                         ઠ
                                                                                                                                            ò
                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                     ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ETSHLELSEADIQQLYGGLLANGDLLAMGDPTLAEDHTIIVEEDKEEESDLEDKQSPTGS 1518
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1519 DSGSVQEDSGSBPKRALPPRKGPNFTWEKFLDPSRPYKCTVCKESFTQ-----KNILLV 1572
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps 110;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBFIESKRPRLELLPDPLLRPSPLLATGQPAGSEDLT----KDRSLTGKLEPVSPPS--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----LSKEELIQNMDRVDREITTMV
                                                                                                              | National Content | National Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content Content 
                                                  GO:0005634; C:nucleus; NAS.
GO:0003700; F:transcription factor activity; TAS.
GO:0030182; P:neuron differentiation; TAS.
GO:0005355; P:regulation of transcription, DNA-dependent; NAS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels 1007;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 3726;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POLY-SER.
W; 915ACBE588A72C98 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (ATYPICAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              tch 2.8%; Score 375.5; DB 1; al Similarity 18.7%; Pred. No. 4.8e-05; 491; Conservative 247; Mismatches 887;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HOMEOBOX 3.
C2H2-TYPE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                    C2H2-TYPE
C2H2-TYPE
C2H2-TYPE
C2H2-TYPE
C2H2-TYPE.
C2H2-TYPE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             C2H2-TYPE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  C2H2-TYPE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TYPE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POLY-GLU.
POLY-ALA.
POLY-ALA.
POLY-GLN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POLY-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POLY-
POLY-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            406567
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2561
2709
2743
3011
3056
3576
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3220
3409
3420
3550
3623
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1863
                            MGD; MGI:99948; Atbfl.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       491
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DNA_BIND
ZN_FING
ZN_FING
DNA_BIND
ZN_FING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DNA BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1459
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ZN_FING
ZN_FING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ઠ
```

	OY 1220 DSAITYRGSITHGTPADVLYKGTITRIIGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLS 1279	1280 YEGGMSVTQCSKEDGRSSSGPHETAAPKRTYDMMEGRVGRAISS	Db 812LPGPKGEPGPSTTGPPGPPGFPGLKGKDG-IPGAPGLEGGRGLP	QY 1339 RHSPHHLKEQHHLRGSITQGIPRSYVEAQEDYLRREAKLLKREGTPPPPPRRDLTEAYK 1398 DD 858	Qy 1399 TQALGPLKLKPAHEGLVATVKEAGRSIHEIPREELRHTPELPLAPR-PLKEGSITQGTP-	Db 879GAKGEPGLPGKEGPQGPPGQPGAPGPRGRGFDGEGLPG	1457LKYDTGASTTGSKKHDV	DD 919 LPGVSGMKGDTGLPGVFGLAGFPGLPGFPGFPGFPGFPFGGPFPGTAKG QY 1501 RACYEESLKSRPGTASSSGGSIARGAPVIVPELGKPRQSPLTYEDHGAPFAGHLP	964	Cy 1556RGSPVTWREPTPRLOEGSLSSSKASQDRKLTSTPREIAKSPHSTVPEHHPHPIS Db 1013 KGDRGFNGLPGEKGEPGPAAROGEKG	Qy 1610 PYEHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYL 1669		OY 1670 IRGYPDTAALENRQTIINDYITSQOMHHNTATAMAQRADMLRGLSPRESSLALNYAAGPR	Qy 1730 GIIDLSQVPHLPVLVPPTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPGGP	1783	: : : : : : : : : :	Qy 1843GGGGSSSRPASHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGIITAVEPSK 1894	1895 PTVLRSTSTSSPVRPAATFPPATHCPLGGTLDGVYPTLMEPVLLPKEAPRVARP	Db 1222 PG-LEGQPGPRGMKGEAGLPGAPGRDGLPGL-PGMKGEAGLPGLPGQPGKS	Qy 1952 -RADTGHAFLAKPPARSGLEPASSPSKGSEPRPLVPPVSGHATIAR	Qy 1997 TPAKNLAPHASPDPPAPASASDPHREKTQSKPFSIQELELRSLGYHGSSYSPEGVEPV	1331 IPGKRGNDGIPGKPGPAGLPGLP	DD 1369 GLPGLPGLPGPPPGQXGBTGPPGQPGLPGCPRGDSGYPGAPGKQG	OY 2107 RPLPESQPSSSPLLOTAPGVKGHQRVVTLAQHISEVITQDYTRHHPQQLS 2156	7747
EMBL; M67507; AAA18014.1; PIR; S16366; S16366.	InterPro; IPR008161; Clg helix. InterPro; IPR008160; Collagen. InterPro; IPR001462; Procollagn4_C. Pfan; PP01413; C4; 2.	Collagen; 25.	<pre>// Procollagnc4; I // Connective tissue:</pre>	icing; Glycoprotein; Signal. 26 POTENTIAL. 1763 COLLAGEN ALPHA 2(IV) CHAIN.	27 42 43 1529	1530 1763 NON 1548 1637 OR	1581 1634 OK 1637 1593 1599 BY SIMIL 1656 1752 OR 1749	1690 1749 OR 1752 (BY SIMILARITY). 1702 1709 BY SIMILARITY. 126 N-LINKED (GLCNAC) (POTENTIAL).	249 O-LINKED (XIL) (IN ISOFORM II) (POI 266 GEQGPRGPQGPPGPVPSTGA	GDIGPAGPPGPPGPREFTGSGSIVGPRGHSGDKGVK isoform II) / FTIG+VSP 001159.	; 168526 MW;	Best Local Similarity 20.7%; Pred. No. 2.9e-05; Matches 400; Conservative 119; Mismatches 673; Indels 742; Gaps 93;	SGNEVPRGECSGPATVNNSSDTESIPSPHTEAAKDTGQNGPK 7	239 SGEAGFRGARGGREGLEGFFGFRGERGLEGGFFGGFFGGRGEAGFFGKLGARGARGEFF 350 769 PPATLGADGPPFGPTPPRRTSRAPIEPTPASEATGAPTPPPAPPSPSAPPPV 821	357 GPPGGGEFSDG-PPGPPGLPGREGQPGPPGADGXPGPPGPQGL 401	822	866 CTEEABEGPAKCKDABAABATABGALKAEKKEGGSGRATTAKSSGAPQDS 915	GP-RGVDGQSIPGLPGKDGRPGLPGRKGEMGLPGVRGPPGDSLNGLPGPP	916	LDLKQLKQRAAAIPPIQVTKVHEPPREDAAPTKPAPPAP 	563 GLPGIPGATGAPGDDGLPGAPGRPGPPGPPGQDGLPGLPGQKGE 606 000 PPPQNLQPESDAPQOPGSSPRGKSRSPAPPADKEAFAAEAQKLPGDPP 1047	PTQLTLRPGPPGYPGQRGETGFPGPRGDEGLPGRPGIVGAPGLPGPPGPRGEP	048 CWISGLPPPVPPREVIKASPHAPDPSAFSYAPPGHP-LPLGLHDTARPVLPRPPT 1101	ISNPPPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEHAKAPVGPVTMGLPLPMDPK	711 AVGPMGPP

```
------GLPGLDGLPGPSG---PPGFAGAKGRDGFPGQPGMPGEKGAP---- 1499
                       2213 SVLGG--GEDGIEPVSPPEGMTEPGHSRSAVYPLLYRDG----EQTEPSRMGSKSPGNTS 2266
                                                ----GLPGFFPGIEGIPGPPGLPGPSGPPGP-SYKDGFLLVKHSQTSEVPQCPPGMVK 1554
                                                                                        SOPGTEIFNMPAITGTGLMTYRSQAVQEHASTNMGLEAIIRKALMGKYDQWEESPPLSAN 2366
                                                                                                                                                                           2367 AFNPLNASASLPAAMPITAADGRSDHTLTSPGGGGKAKVSGRPSSRKAKSPAPGLASGDR 2426
                                                                                                                                                                                                   1659 GWNSLWIGYSP--AM------HTGAGAEGGGS------LSSPGSCLEDFRA 1696
                                                                                                                                          | | : | : | : | ST--TAPIPMMPVSEGGIEPYISRCAVCEAPAN-----VIAVHSQTIQIPN-CPN
                                                                        2267 QPPAFFSKLTESNSAMVKSKKQ--------EINKKLNTHNRNEPEYNI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TISSUE=Brain;
MEDINE=SG66099; PubMed=7485154;
MEDINES-G66099; PubMed=7485154;
Onodera O., Oyake M., Takano H., Ikeuchi T., Igarashi S., Tsuji S.;
"Molecular cloning of a full-length cDNA for dentatorubral-
"Molecular cloning of a full-length cDNA for dentatorubral-
pallidoluysian atrophy and regional expressions of the expand alleles
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                "Structure and expression of the gene responsible for the triplet repeat disorder, dentatorubral and pallidoluysian atrophy (DRPLA)."; Nat. Genet. 8:177-182(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE=97228904; PubMed=9074930; Ansari-Lari M.A., Shen Y., Muzny D.M., Lee W., Gibbs R.A.; Ansari-Lari M.A., Shen Y., Muzny D.M., Lee W., Gibbs R.A.; Large-scale sequencing in human chromosome 12p13: experimental and computational gene structure determination."; Genome Res. 7:268-280(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
TISSUE=Brain, and Cerebellum;
MEDLINE=95144175; PubMed=7842016;
Nagafuchi S., Yanagisawa H., Ohsaki B., Shirayama T., Tadokoro K.,
Inoue T., Yamada M.;
                                                                                                                                                                                                                                                                                                                                                                                                      Homo sapiens (Human).
Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia, Eutheria, Primates, Catarrhini, Hominidae, Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
MEDLINE=96262314; PubMed=8965642;
Margolis R.L., Li S.-H., Young W.S., Wagster M.V., Stine O.C.,
Kidwai A.S., Ashworth R.G., Ross C.A.;
"DRPLA gene (atrophin-1) sequence and mRNA expression in human
brain.";
                                                                                                                                                                                                                                                                                                                            P54259; Q99495; Q99621; Q9UEK7; COCT-1996 (Rel. 34, Created) (10-OCT-2003 (Rel. 42, Last sequence update) (10-OCT-2003 (Rel. 42, Last annotation update) Atrophin-1 (Dentatorubral-pallidoluysian atrophy protein).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A., AND POLYMORPHISM OF POLY-GLN REGION
                                                                                                                                                                                                                                                                                                                    PRT; 1185 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Brain Res. Mol. Brain Res. 36:219-226(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           in the CNS.";
Am. J. Hum. Genet. 57:1050-1060(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE OF 1-76 FROM N.A.
MEDLINE=97005364; PubMed=8852663;
                                                                                                                                                                                                                             2427 PPSVSSVHSEGDCN 2440
                                                                                                                                                                                                                                                   1697 TPFIECNGARGTCH 1710
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TISSUE=Brain;
                                                                                                                                                                                                                                                                                                                    HUMAN
                                                1500
                                                                                                 1555
                                                                                                                       2307
                                                                                                                                                  1612
1461
                                                                                              g
                                                                                                                                               g
                                                                                                                                                                                               g
                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                             ò
                                           임
                                                                       à
                                                                                                                       ò
                                                                                                                                                                         ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isp-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DISEASE: Defects in DRPLA are the cause of dentatorubral-pallidoluyaian atrophy (DRPLA) [MIN:125.70], an autosomal dominant neurodegenerative disorder characterized by a loss of neurons in the dentate nucleus, rubram, glogus pallidus and Luya'body. Clinical features are myoclonus epilepsy, dementia, and cerebellar ataxia. Onset of the disease occurs usually in the second decade of life and death in the fourth.

CAUTION: Ref.2 sequence differs from that shown due to several
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -1- POLYMORPHISM: The poly-Gln region of DRPLA is highly polymorphic (7 to 23 repeats) in the normal population and is expanded to about 49-75 repeats in DRPLA parients. Longer expansions result in earlier onset and more severe clinical manifestations of the
Yanagisawa H., Fujii K., Nagafuchi S., Nakahori Y., Nakagome Y., Akane A., Nakamura M., Sano A., Komure O., Kondo I., Jin D.K., Soerensen S.A., Potter N.T., Young S.R., Nakamura K., Nukina N., Nagao Y., Tadokoro K., Okuyama T., Miyashita T., Inoue T., Kanazawa I., Yamada M., Muique origin and multistep process for the generation of expanded by Liplet repeats."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Wood J.D., Yuan J., Margolis R.L., Colomer V., Duan K., Kushi J., Kaminsky Z., Kleiderlein J.J. Jr., Sharp A.H., Ross C.A.;
"Atrophin-1, the DRPLA gene product, interacts with two families of WW domain-containing product, interacts with two families of Loll. Neurosci. 11:149-160(1998).
-!- SUBUNIT: Interacts with WWPl and WWP2.
-!- TISSUB SPECIFICITY: Relatively high levels in the brain, ovary, testis and prostate. Lower levels in the liver, thymus and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       C.A.;
                                                                                                                                                                                                                                                                                                                                                                                           TISSUE=Brain cortex;
MEDLINE=93315145; PubMed=8325528;
Li S.-H., McInnis M.G., Margolis R.L., Antonarakis S.E., Ross C.
"Novel triplet repeat containing genes in human brain: cloning,
expression, and length polymorphisms.";
[7]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MIM; 125370; -
GO; GO:0005737; C:cytoplasm; TAS.
GO; GO:0005634; C:nucleus; TAS.
GO; GO:0005614; C:nucleus; TAS.
GO; GO:000517; F:protein binding; IPI.
GO; GO:0007417; P:central nervous system development; TAS.
PIGERP: PF03154; Atrophin.
PFM; PF03154; Atrophin.1; 2.
PRINTS; PR01222; ATROPHIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Polymorphism; Epilepsy.
SER/GLU-RICH (MIXED CHARGE).
POLY-PRO.
POLY-SER.
POLY-SER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; D31840; BAA06626.1; -.
EMBL; D38529; BAA07534.1; ALT_FRAME.
EMBL; U23851; AAB50276.1; -.
EMBL; U4924; AAB51321.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL, D63808; BAA23631.1; -.
EMBL, L10377; -; NOT_ANNOTATED_CDS.
PIR; G01763; G01763.
Genew, HGNC:3033; DRPLA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE=98313405; PubMed=9647693;
                                                                                                                                                                                                                                                                               Hum. Mol. Genet. 5:373-379(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       INTERACTION WITH WWP1 AND WWP2.
                                                                                                                                                                                                                                                                                                                                                            SEQUENCE OF 470-725 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  expansion;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         305
305
382
397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             frameshifts.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Genew; HGNC:3033
MIM; 607462; -.
MIM; 125370; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              repeat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               leukocytes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               disease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
DOMAIN
DOMAIN
```

	1079ASGSHLTRIPYPAGTLENPLLEHE 2156 SAPLPAPLYSFPGASCPVLDLRRPPSDI 1126 SAPMSA-AHQLQAMHAQSAELQRLALEQQQMI T 45 HUMAN	utele Homo.	RA Willis T.G., Zalcberg I.R., Coignet L.J.A., Wlodarska I., Stul M., RA Jadayel D.M., Bastard C., Treleaven J.G., Catovsky D., Silva M.L.M., RA Dyer M.J.S.; RT "Molecular cloning of translocation t(1;14) (q21;q32) defines a novel
Gaps APPPV PAKGK DE DE DE DE DE	ARY DEPTITION FOR THE STANKES DESCRIBED TO THE STANKES OF THE STAN	394 SSASPFPASQALPSYPHSFPPPTSLSVSNQPFKYTQPSLFSQAV 437 1269 IYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETAAPKRTYDMMEGRVGRAISSASIE 1328 438GILLANSNAH 458 1329GILGARAIPPERHSPHHLKEQHHIRGSITCGIPRSYVEAQEDYLRREAKLIKR 1380 459 PGPFPPSTGAQSTAHPPVSTHHH	: GAP-Y QSPLT

```
1884 KGIITAVEPSKPTVLRSTSTSSPVRPAATFPPATHCPLGGTLDGVYPTLMEPVLLPK-EA 1942
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PLTMASP----AMLGNVESGGPPPPTASQPASVNIPGSLPSSTPYTMPPEPTLSQNPLSI 1008
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ----AKNLAP-----HH-----ASPDPPAPASASDPHREKTQSKPFSIQELBLRSLG 2042
                                                                                                                                                                                                                                                                                                                                                               1206 GSITKGI------PSTRVP----SDSAITYRGSITHGTPADVLYKGTITRIIG 1248
                                                                                                                                                                   PVLPRPPTISNPPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEHAKAPVGPVTMGLP 1153
                                                                                                                                                                                                                                                                                                                                                                                                                                                        EDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGR-SSSGPPHETAAP 1307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1365 EAQEDYLRREAKLLKREGTPPPPPPSRDLTEAYKTQALGPLKLKPAHEGLVATVKEAGRS 1424
- KPLPQQPPVPANQDQNSSQNTRLQPTPP1PAPAPKPAAPPRPLDRESPGVENKL1PSVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1597 HSTVPEHHPHPISPYEHLLRGV--SGVDLYRSHIPLAFDPTSIPRGIP---LDAAAAYYL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1652 PRHLAPNPTYPHLYPPYLIRGYPDTAALENRQTIINDYITSQQMHHNTATAMAQRADMLR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ----ISVAGSQVH
                                                                                                                                                                                                                       -----RMLFP
                                                                                                                                                                                                                                                                   LPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGV---PTAQ-----EASVLRGTALGSVPG
                                                                                                                                                                                                                                                                                                             ---DEKE---FICAQ-----SGGPQONPGVLDGPQKKPEGPIQAMMAQSQSLGKGPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1308 KRTYDM--MEGRVGRAISSASIEGL-MGRAIPPERHSPHHLKEQHHIRGSITQGIPRSYV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ----PNMPGSQ-----MRL-PGFAGMINSEMEG---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ------PNVPNPASRPGLS-GVSWPDDVPKIPDGRNFPPGRGIFSGPGRGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1483 TFPPVHPLDVMADARALERACYEESLKSR-----PGTASSSGGSIARGAPVIVPELGKPR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1538 QSPLTYEDHGAPFAGHL-PRGSPVTMREPTPRLQEGSLSSSKASQDRKLTSTPREIAKSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ----SDMLPAQO------KMVPLPF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1712 GLSPRESSLALNYAAGPRGIIDLSQVP----HLPVLVPPTPGTPATAMDRLAYLPTAPQP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GEHPQQ-----EYGMGPRPFLPMSQGPGSNSGLRNLREPI-GPDQRTNSRLSHMPPLPLN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1768 FSSRHSSSPLSPGGPTHLTKPTTTSSSERERDRDRERDRDREREKSILTSTTTVEHAPIW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  857 SPGI------NPLKSPTMHQVQSPMLGSPSGNLKSPQTPSQL--AGML
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            P-RVARPERPRADTGHAFLAKPPARSGLEPASSPSKGSEPRPLVPPVSGHATIARTP---
                                                                          REVIKASPHAPDPSAFSYAPPGHPLPLGLHDTAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ------tklooefyeekr----rkpequvyoocslobmvhohgprgvvrgp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -----IRPSMEMNR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1828 RPGTEQSSGSSGSSGGGGSSSRPASHAHQ-HSPI--SPRTQDALQQRPSVLHNTGM-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             897 AĞPAAAASIKSPPVLGSAAAS----PVHLKSPSLPAPSPGWTSSPKPPLQSPGIPPNHKA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1425 IHEIPREELRHTPELPLAPRPLKEGSITQGTPLKYDTGASTTGSKKHDVRSLIGSPGR
                                                                                                                                                                                                                                                                                                                                                                                                           422 PRIDVGAPFGPQGHRDVPFSPDEMVPPSMNSQSGTIGPDHLDHMTPEQIAW----
                                                                                                                      SPASST-----PLPPDGTGPNSTPNNRAVTPVSQGSNSSSADPKAPPPVSSG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PPPYQMTPSEGWAPGGTEPFS-DGINMPHSLPPRGMAPH-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      824 PSSNPTSLNTAP-----PVQRGLGRKPLD------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ----EPPTLGENPDGLSQEQ----LEHRERSLQTLRDIQ-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ..
                                                                          DPPCWTSGLPFPVPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          620 RFP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1943
                                                                                                                                                                                                                                                                                                                                                                                                                                                             1249
                                                                          1045
                                                                                                                                                                                                                    340
                                                                                                                                                                                                                                                                      1154
                                                                                                                                                                                                                                                                                                                  376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     515
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 576
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               744
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           770
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              953
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     553
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Вb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           유
                                                                                                                    g
                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                       Š
                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMED outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                992 TKPAPPAPPPPQNLQPESD-----APQQPGSSPRGKSRSPAPPADKEAFAAEAQKLP--G 1044
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ----GSMGLKNGAGNGAKGKGKRERSISADSFDQRDPG-TPNDDSDIKECNSADHIKSQD 135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CAUTION: It is uncertain whether Met-1 or Met-27 is the initiator. CAUTION: Ref.1 sequence differs from that shown due to a frameshift in position 1391.
                                                                                                                           ŝ
                                                                                                                                                                                                                                                                                                                                                                                DISBABE: Involved in a t(1,14) (q21,q32) chromosomal translocation found in a patient with precusor B-cell acute lymphobiastic leukemia (ALL). This translocation leaves the coding region intact, but may have pathogenic effects due to alterations in the expression level of BCL9. Several cases of translocations within
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               724 LHASGNEVPRGECSGPATVNNSSDTESIPSPHTEAAKDTGONGPKPPATLGADGPPPGPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MHSSNPKVR----SGP-----SGNTQSSPKSKQEVM-----VRPPTVMS----PSGNP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TPPRRTSR-----APIEPTPASEATGAPTPPPAPPSPSAPPVVPKEEKEETAAAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PVEEGEEQKPPAAEELAVDTGKAEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         KEGGSGRAT----TAKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPT---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----EPTPAQKTPAKVV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ---GDPRANASPQKPL------DLKQLK-QRAAAIPPIQVTKVHEPPREDAAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 YVFSTEMANKAAEAVLKGQVETIVSFHIQNISNNKTERSTAPLNTQISALRNDP----
                                                                                                                                                                                                                                             Wnt pathway
                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                       testis, ovary and small intestine, and at lower levels in spleen, colon and blood.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                           FUNCTION: Involved in signal transduction through the Wnt pathw SUBUNIT: Binds to beta-catenin (CTNNB1), PYGO1 and PYGO2.
SUBCELLULAR LOCATION: Nuclear (Probable).
TISSUE SPECIFICITY: Detected at low levels in thymus, prostate,
                                                                                                                         Chatterjee
                                                                                                                                                                        recruitment
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            the 3' untranslated region of BCL9 have been found in B-cell
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               al Similarity 19.7%; Pred. No. 4.6e-05; 378; Conservative 195; Mismatches 686; Indels 658;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 1426;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   translocation; Proto-oncogene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POLY-ALA.
POLY-PRO 3.
MW; A240A487716B7F1B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          QLDSKFSNQGKQGGSASQSQPSPCDSKSGGHT-PKALPGPG-----
                                                                                                                      Kramps T., Peter O., Brunner E., Nellen D., Froesch B., Murone M., Zuellig S., Basler K.;
                                                                                                                                                                   "Mnt/wingless signaling requires BCL9/legless-mediated pygopus to the nuclear beta-catenin-TCF complex."; Cell 109:47-60(2002).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SQHTPHSMTPSNATAPRSSTPSHGQTTAT------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CTNNB1-BINDING.
POLY-PRO 1.
POLY-PRO 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 364;
Pred. No. 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRO-RICH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; Y13620; CAA73942.1; ALT_FRAME.
Genew; HGNC:1008; BCL9.
                                                                                                  MEDLINE=21952490; PubMed=11955446;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MIM; 602597; ...
Nuclear protein; Chromosomal 1
Wht signaling pathway.
DOMAIN 347 377 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          149314
                           91:1873-1881 (1998)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ą,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     malignancies
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 970
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 н
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  784
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              936
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         968
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       949
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local
       gene (1
Blood (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Loc
Matches
```

1537

575

1711

8 셤 ð g 2 8

중 점

g ⋩

ð

```
. අ
.
     SFFF
                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                        ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 윰
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                        PAFFSKLTESNSAMVKSKKQEINKKLNTHNRNEPEYNISQPGTEIFNMPAITGTGLMTYR 2328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1371
MMSRMSKFAMPSSTPLYHDAIKTVASSDDDSPPARS--PNLPSMNNMP------GMG 1057
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----BOAPRMGL-----ALPGM-----GGPGPVGTPDIPLGTAPSMPGHNPMRPPAF 1321
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2385 AADGR--SDHTLTSPGGGGKAKVSGRPSSRKAKSPAPGLASG-DRPPSVSSVHSEGDCNR 2441
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RIPLINRVWEDRPSSAGSIPFPYNPLIMRLQAGVMAS-----PPGLPAGSG 2489
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ----PGPVGS-----PGMMMSMQGMMGPQQNIMIPPQMRPRGMAADVG 1410
                                                                           ---GHGSQEPPMVPQG
                                                                                                                                                                                                                                                                                    1322 LOOGMMGPHHRMMSP---AQSTMPGQPTLMSNPAAAVGMIPGKDRGPAGLYTH-----
                                                 YHGSSYSPEGVEPVSPVSSPSLTHDKGLPKHLEELDKSHLEGELRPKQPGPVKLGGEAAH
                                                                                                                                                    LPHLRPLPESQPSSSPLLQTAPGVKGHQRVVTLAQHISEVITQDYTRHHPQQLSAPLPAP
                                                                                                                                                                                                                                                  2163 LYSPPGASCPVLDLRRPPSDLYLPPPDHGAPARGSPHSBGG------KRSPEPN
                                                                                                                                                                                                                                                                                                                                                     2211 KTSVL - GGGEDGIEPVSPPEGMTEPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EPDLSRIIPSEK---PSQTLQYPPRGEVPGRKQPQ----GPGPGFSHMQGMMG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SQAVQEHASTNMGLEAIIRKALMGKYDQWEESPPLSANAFNPLNASASLPAAMPIT----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   brain,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE-20175430; PubMed=10708517;
Smith J.S., Tachibana I., Pohl U., Lee H.K., Thanarajasingam U.,
Portier B.P., Ueki K., Billings S., Ramaswamy S., Mohrenweiser H.W.,
Schaithauer B.W., Louis D.N., Jenkins R.B.;
"A transcript map of the chromosome 19q-Arm glioma tumor suppressor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Genomics 64:44-50(2000).
-!- TISSUE SPECIFICITY: Expressed at moderate levels in heart, brair
placenta, skeletal muscle, and pancreas, and at lower levels in
lung, liver, and kidney.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Craniata, Vertebrata; Euteleostomi;
Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Glioma tumor suppressor candidate region gene 1 protein.
GLYSCR1.
Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Æ
                                                                                                                                                                                                      ---- BNI PPHGVPM---GPGLMSHNPIM---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1509
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POLY-GLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (Rel. 40, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; AF182077; AAF62874.1; -. Genew; HGNC:4332; GLTSCR1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Bukaryota; Metazoa; Chordata;
Mammalia; Butheria; Primates;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                37.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MIM; 605690;
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             16-OCT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HUMAN
                                                 2043
                                                                                                                                                  2103
                                                                                                                                                                                                                                                                                                                                                                                                                                                        2269
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2329
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2442
                                                                                                                                                                                                      1100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 region."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q9NZM4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GSR1 HUMAN
                                                                                                   셤
                                                 ò
                                                                                                                                               ò
                                                                                                                                                                                           요
                                                                                                                                                                                                                                                  ሯ
                                                                                                                                                                                                                                                                                       요
                                                                                                                                                                                                                                                                                                                                                   ઠે
                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PRESENTATION OF THE PR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         용
```

```
995 APPAPP-PPQNLQPESDAPQQP-----GSSPRGKSRSPAP-PADKEAFAAEAQ- 1040
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1352 RGSITQGIPRSYVEAQEDY-----LRREAKL------LKREGTPPPPPPPSKDLTEAYK 1398
                                                                                                                                                                                                                                         51,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                --KLPGDPPCW----TSGLPFPVPPRE---VIKASPHAPDPSA----FSYAPPGHPLPLG 1087
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1088 LHDTARPVLPRPPIISNPPPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEHAKAPVGP 1147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1148 VTMGLP--LPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEASVLRGTALGSVPG 1205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1206 GSITKGIPSTRVPSDSAI-----TYRGSITH-----GTPA--DVL-----YKGTI- 1243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -----prilopskeacplehikhggsvlhpdyktappspedalhrllpyhvyggalp 1074
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----KTY-EARSRIGLKLKIKQEAGLSKVVHNTALDPVHQPPPPPATLKVAEP-- 1294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ---LIRELA 1382
                                                                                                                                                                                                                                                                                                                                                                                                             845
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      631
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               889
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     691
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ALKAEKKEGGSGRATTAKSSGAPQDSDSSA-----TCSADEVDEAEGGDKNRLLSPR 941
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PSL--LTPTGDPRANASP--OKPLDLKQLKQRAAAIPPIQVTKVH---EPPREDAAPTKP 994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      865 PHKSPTPPPTLHLVPEPAAPPPPPRTFOMVTTPFPALPOPKALLERFHQVPSG----II 920
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    921 LONKAGGAPAAPQTSTSLGPLTSPA--ASVL----VSGQAPSGTPTAP--SHAPAPAP 971
                                                                                                                                                                                                                                                                                                                                                     SLPTQSQPAPAGPAATTVLQGVTLPPSAVAMLNTPDGLVQPATPAAATGEAAPVLTVQPA
                                                                                                                                                                                                                                                                                               SSDTESIPSPHTEAAKDTGONGPKPPATLGADGPPPG---PPTPPRRTSRA----PIEPT
                                                                                                                                                                                                                                                                                                                                                                                                             P-ASEATGAPTP-----PPAPPSP-SAPPPVVPKEEKEETAAAPPVEEGEEQKP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PAAEELAVDTG---------KAEEPVKSECTEEAEEGPAKGKDAEAAEATAEG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PSATPTAILTQDSLQMFLPQERSQQPLSAEGPHLSVPASVIVSAPPPAQDPAPATPVAKG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  805 GPPQPPLRPQSQPPEGPLPPAPHLPPSSTSSAVASSSETSSRLPAPTPSDFQLQFPPSQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----ETAAPKRIYDMMEGRVGRAISSASIEGLMGRAIPPERHSPHHLKEQHHI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             --APRPLKEGSITQGT--PLKYDTGASTTGSKKHDVRSLIGSPGRTFPPVHPLDVMADAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POAPPAVSTPLPLGLQQPQAQQPPQAPTPQAAAPPQATTPQPSPGLASSP--EKIVLGQP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1075 SPSDYHKVDEEFETVSTOLLKRTQAMLNKYRLLLLEESRRVSPSAEMVMIDRMFIQEEKT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            --RLDRGREDSLPKGHV----IYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ----PPRPPPPPP---PTGQMNGTVDHPPPAAPERKPLGTAPHCPRLPLRKTYRENVGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   972 AATGLPPLLPAENKAFA--SNLPTLNVAKAASSGPGKPSGLQYESKLSGLKK-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1191 LPTKLVIRHGGAGGSPSVTW----ARASSSLSSSSSSAASSLDADEDGPMPSRNRPPI
                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                          Query Match 2.7%; Score 360; DB 1; Length 1509; Best Local Similarity 23.0%; Pred. No. 6.4e-05; Atches 249; Conservative 104; Mismatches 420; Indels 308;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TOALGPLKLKPAHEGLVATVKEAGRSIHEIPREELR-----HTPELPL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ----SPS----SPS-----
  POLY-PRO.
POLY-SER.
POLY-PRO.
POLY-PRO.
WW: 7C5144F443CE6821 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1347 PGAPEGTPAGRARGGSPAPLPAKVDEATSG------
884 889
1214 1225
1282 1286
1294 1304
1509 AA; 152991 M
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               942
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      574
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     632
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              890
                                                                                                                                                                                                                                                                                            745
                                                                                                                                                                                                                                                                                                                                                                                                             798
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               846
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       692
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1247
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1442
                                                                                                                                                                             Query Match
                                                                                                                       SEQUENCE
     DOMAIN
DOMAIN
DOMAIN
                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
```

```
Mac
Local S.
323;
                                                                                                                                                                                                                                              224
                                                                                                                                                                                                                                                                                             249
                                                                                                                                                                                                                                                                                                                   549
                                                                                                                                                                                                                                                                                                                                          306
                                                                                                                                                                                                                                                                                                                                                                                       359
                                                                                                                                                                                                                                                                                                                                                                                                                                                            869
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         738
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                531
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             591
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 885
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1040
                                  SEQUENCE
                                                        Query Match
         DOMAIN
DOMAIN
                                                                             Matches
FF
                                                                                                                       д
                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                        셤
                                                                                                                                                                                               ò
                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                  엄
                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                     ઠે
                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        용
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                           AVEDELYQRMLKGPPPEPAASAAQGTGDPDWEAPGLPPAKRRKSESP---DVDQASFSSD 1439
        ALERACYEESLKSRP----GTASSSGGSIARGAPVIVPELGKPROSPLTYEDHGAPFAGH 1553
                                                                             1440 SPQDD--TLTEHLQSAIDSILNLQQAPGRTPAPSYPHAASAGTPASPPLHRPEAYPPSS 1497
                                                       1554 LPRGSPVTMREPTPRLQEGSLSSSKASQDRKLTSTPREI-AKSPHSTVPEHHPHPISPYE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DEV. BIOL. 205:309-321(1999).
-!- TISSUE SPECIFICITY: CNS STEM CELLS.
-!- DEVELOPMENTAL STAGE: UPON TERMINAL NEURAL DIFFERENTIATION, NESTIN IS DOWN-REGULATED AND REPLACED BY NEUROFILAMENTS.
-!- SIMILARITY: Belongs to the intermediate filament family.
                                                                                                                                                                                                                                                                              Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Buteleostomi,
Mammalia, Butheria, Primates, Catarrhini, Hominidae, Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE OF 296-309 FROM N.A.
MEDILIBE-9911735, PubMed=9917366;
Yaworsky P.J., Kappen C.;
"Heterogeneity of neural progenitor cells revealed by enhancers
                                                                                                                                                                                                                                                                                                                                                              Dahlstrand J., McKay R.D.G., Zimmerman L.B., Lendahl U.;
Characterization of the human nestin gene reveals a close
evolutionary relationship to neurofilaments.";
J. Cell Sci. 103:589-597(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MIM; 600915; -.
GO; GO:0005882; C:intermediate filament; NAS.
GO; GO:0007417; P:central nervous system development; NAS.
InterPro; IPR001664; IF.
PRAM; PF00038; filament; 1.
PROSITE; PS00226; IF; 1.
Intermediate filament; Coiled coil; Neurone.
                                                                                                                                                                                                         Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                    PRT; 1618 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                coil; Neurone
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TAIL.
COIL 1A.
LINKER 1.
COIL 1B.
LINKER 12.
                                                                                                                                                                                                                                                                                                                                        TISSUE=Placenta;
MEDLINE=93123384; PubMed=1478958;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; AF004335; AAB64426.1; -. PIR; S21424; S21424. Genew; HGNC:7756; NES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; X65964; CAA46780.1;
                                                                                                                                                                                    STANDARD;
                                                                                                                                                                                                   01-FEB-1996 (Rel. 33, C:
01-FEB-1996 (Rel. 33, Le
28-FEB-2003 (Rel. 41, La
Mettin.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               43
55
151
172
                                                                                                                                                                                                                                                                     sapiens (Human).
                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       the nestin gene.";
                                                                                                                                                                                                                                                                                                      NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   313
8
44
56
152
                                                                                                                                                                                              P48681; O00552;
01-FEB-1996 (Re]
                                                                                                    H 1613
                                                                                                                         H 1498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MIM; 600915; -
                                                                                                                                                                                     HUMAN
                               1383
                                                                                                                          1498
                                                                                                    1613
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                        HUMAN
                                                                                                                                                                                                                                                                     Ношо
                                                                                                                                                                        NEST
                             셤
                                                      ð
                                                                          a
                                                                                                  à
                                                                                                                        셤
```

```
147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ----DPEGOSOQVGTPGLOAPQGLPEAIEPLVEDDVAPGGDQASPEVMLGSE---PAMGE 1092
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        148 VSPPSPPHTDPELELVPPRLSKEELIQNMDRVDREITMVEQQISKL-----KKKQQQLE 201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         : ||:|
389 I-PPTPQAPSP------AVDAEIRAQDAPLSLLQTQGGRKQAPEPLR 428
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----RSLVQIIYDE---NRKKAEAAH-----RILEGL 248
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GPQVE---LPLYNQPSDTRQYHENIKINQAMRKKLILYFKRRNHARKQWKQKFCQRYDQL 305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | ::|||::| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:|
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           791 RAPIEPTPASEATGAPTPPPAPPSPSAPPPVVPKEEKEEETAAAPPVEEGEEQKPPAAEE 850
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEALEKKVERI -----ENNPRRRAKESKVREYYEKQFPEIRKQRELQERMQSRVGQRG 358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           419 PMKVYKDROVMNWWSEQEKETFR -- EKFMOHPKNFGLIASFLERKTVAECVLYYYLTKKN 476
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  477 ENYKSL-----VRRSYRRRGKSQQQQQQQQQQQQQQQPMPRSSQEEKDEKEKEAE 530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          KEEEKPEVENDKEDLLKEKTDDTSGEDNDEKEAVASKGRKTANSQGRRKGRITRSMANEA 590
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----LLEH 630
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        631 GRNWSAIARMVGSKTVSQC-----KNPYFNYKKRQNLDEILQQHKLKMEKERNARRKK 683
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                684 KKAPAAASEEAAFPPVVEDEEME-----ASGVSGNEEEMVEEAEALHASGNE--VPRGE 735
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       736 CSGPATVNNSSDTESIPSP---HTEAAKDTGQNGPKPPATL--GADGPPPGPPTPPRRTS 790
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LAV-----EAEEGPAKGKDAEATET 886
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ENHKSLRSLBEGDOGFTLRTLEKETÓGRRRSLGEGDÓMTLRP----PEKVDLEPLKSLD 791
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   835 NLETLKSPETQAPL------WTPEEINKSGGNESSRKGNSRTTGVCGSEPRDIQTP 884
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         329 KLELQFPRTPEGRRLGSLLPVLSPTSLPSPLPATLETPVPAFLKNOEFLQARTPTLASTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 97 KSEMEFIESKRPRL--ELLP--DPLLRPSPL---LATGOPA--GSEDLTKDRSLTGKLEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SGLSMSAARSEHEVSEIIDGLSEQENLEKQMRQLAVIPPMLYDADQQRIKFINMNGLMAD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PLRSLED------ENKEAFRSLEKENQEP-----LKTLEEEDQSIVRPLET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                  502;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NSEEAITPQQSAELASMELNESSRWTEEEM-----ETAKKG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                202 EEA--AKP----PEPEKP-----VSP-----PPIESKH----
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Conservative 196; Mismatches 511; Indels
173 191 COLL 2A.
192 194 LINKER 2.
195 312 COLL 2B.
1618 AA; 176704 MW; C9E9AA48CG6534D0 CRC64;
                                                                                                                                                                                                                                                                                                                          1 2.7%; Score 358; DB 1; Similarity 21.1%; Pred. No. 7.9e-05;
```

```
340 IRKQRELQERMQSRVGQRGSGLSMSAARSEHEVSEIIDGLSEQENLEKQMRQLAVIPPML 399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MUTAGEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      772
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  259
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   962
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sest Local
                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                             REPEAT
                                                                                                                                                                                                                                                                                                                                                        REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                                                                Repeat
                                                                                                                                                                                                                                                                                                                                 REPEAT
                                                                                                                                                                                                                                                                                                                                                                     REPEAT
                                                                                                                                                                                                                                                                                                                                                                                 REPEAT
                                                                                                                                                                                                                                                                                                                                                                                              REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                          REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches
 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                  1058
                                                                                                                                                                                                                                                                              1443
                                                                          1232
                                                                                                                           1274
                                                                                                                                                    1114
                                                                                                                                                                                                    1161
                                                                                                                                                                                                                                                                                                                               1502
                                                                                                                                                                                                                                                                                                                                                                                1553
                                                                                                                                                                                                                                                                                                                                                                                                                                  1590
                                                                                                                                                                                                                    |-||-||
PAVLASEGLEEPSEKEEGEEGEERCGRDSDLSEEFEDLGTEAPF-----LPGVPGEV 1383
                                                                                                                                                                                                                                                      1185
                                                                                                                                                                                                                                                                                                     --VPTAQEASVLRGTALG-SVP-----GGSITKGIPSTRVPSDS--AITYRGSITHGTPA 1235
                                                                                                                                                                                                                                                                                                                                                       DVLYKGTITRIIGEDSPSRLDRGREDSLPKGHVI-----YEGKKGHVLSYEGG-MSV 1286
                                                                                                                                                                                                                                                                                                                                                                                                         1287 TQCSKEDG------RSSSGPPHETAAPKRTYDMMEGRVGRAISSASIEGLMGRAIP 1336
 940
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAINESPIRAGE Dawley, and Fischer 344/N;
MEDLINE=95148647; PubMed=7846077;
Kakiuchi H., Watanabe M., Ushijima T., Toyota M., Imai K.,
Watanabe M., Ushijima T., Toyota M., Imai K.,
Weisburger J.H., Sugimura T., Nagao M.;
"Specific 5'-GGGA-3'-->5'-GGA-3' mutation of the Apc gene in rat colon tumors induced by 2-anino-1-methyl-6-phenylimidazo[4,5-b]pyridine.";
Proc. Natl. Acad. Sci. U.S.A. 92:910-914(1995).
-!-FUNCTION: Tumor suppressor. Procees rapid degradation of CTNNB1 and participates in Wnt signaling. APC activity is correlated with
               --GGLGTEFSELPGKSRDPWEPPREGREESEARAPRGAEEAFPAETLGHTGSD----AP
887 AEGALKAEKKE-GGSGRATTAKSSGAPQDSDSSATCSADEVDEAE----GGDKNRLLSP
                                                 941 RPSLLTPTGDPRANASPQKPLDLKQLKQRAAAIPPIQV--TKVHEPPREDAAPTKPAPPA
                                                                         -----VPPVLVSPSPTYTPILEDA-----
                                                                                                  PPPPQNLQPESDAPQQPGSSPRGKSRSPAPPADKEAFAAEAQKLPGDPPCWTSGLPFPVP
                                                                                                                         1233 ----PGLOPQAEGSQEASWGVQGRAEAGKVESEQEELGS-----GEIP---EGL----
                                                                                                                                                  PREVIKASPHAPDPSAFSYAPPGHPLPLGLHDTARPVLPR-PPTISNPPLISSAKH---
                                                                                                                                                                          1275 OEEGEESREESEEDELGETLPDSTPLGFYLRS---PTSPRWTPLESRGHPLKETGKEGWD
                                                                                                                                                                                                    PSVL------ERQIGAISQGMSVQLHVPYSE-HAKAPVGPVTMGLPLPMDPKKL
                                                                                                                                                                                                                                                     ------RGOAGPPESLG--
                                                                                                                                                                                                                                                                            1384 AEPLGQVPQLLLDPAAWDRDGESDGFADEESGGEEGEEDQEEGREPGAGRWGPGSSVGSL
                                                                                                                                                                                                                                                                                                                             1444 QALSSSQRGEFLESDSVSVSVPWDDSLRGAVA-GAPKTALETESQDSAEPSGSEEESDPV
                                                                                                                                                                                                                                                                                                                                                                               SLEREDKV-----PGPLEIPSGMEDAGPGADIIGVNGQGPNLEGKSQHV---NGGVMNG
                                                                                                                                                                                                                                                                                                                                                                                                                                  ----SSAG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Craniata; Vertebrata; Euteleostomi;
Sciurognathi; Muridae; Murinae; Rattus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              "cDNA cloning of the rat APC gene and assignment to chromosome 18."i Mamm. Genome 6:746-748(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      H., Watanabe M., Imai K., Yachi A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Adenomatous polyposis coli protein (APC protein)
                                                                                                                                                                                                                                                                                                                                                                                                                                1554 LEQSEESGARNALVSEGDRGSPFQEEEGSALKR----
                                                                                                                                                                                                                                                                                                                                                                                                                                                        1337 PERHSPHHLKEQHHIRGSITQGIPRSYVEAQE 1368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2842 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (Rel. 40, Created)
(Rel. 40, Last sequence update)
(Rel. 42, Last annotation update
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN=Fischer 344/N; TISSUE=Brain;
MEDLINE=96116966; PubMed=8563176;
                                                                         1201 SP---WPLGSEEAEED-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Toyota M., Ushijima T., Kakiuchi
Sugimura T., Nagao M.;
                                                                                                                                                                                                                                                   A-PFSGVKQEQLSP-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Bukaryota, Metazoa, Chordata,
Mammalia, Eutheria, Rodentia,
NCBI_TaxID=10116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Rattus norvegicus (Rat)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      16-OCT-2001
16-OCT-2001
10-OCT-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MUTAGENESIS
                                                                                                  666
                        1148
                                                                                                                                                  1059
                                                                                                                                                                                                                            1332
                                                                                                                                                                                                                                                     1162
                                                                                                                                                                                                                                                                                                                                                       1236
                                                                                                                                                                                                    1115
                                                                                                                                                                                                                                                                                                     1186
                                                                                                                                                                                                                                                                                                                                                                                1503
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1591
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APC RAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 48
                       셤
                                                δ
                                                                    g
                                                                                                 δ
                                                                                                                        염
                                                                                                                                                                   셤
                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                  Š
                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                            용
                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                  ઠે
                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial
                                                                                                                                                                                                                             entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           127;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     942 NRTCSMPYAKVEYKRSSNDSLNSVTSSDGYGKRGQMKPSVESYSEDDEGKFCSYGQYPAD 1001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           667
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 713
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         830
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               881
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ...-----QPSDTRQYHENIKINQAMRKKLILYFKRRNH--ARKQWK 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  941
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QKFCQ-----RYDQLMEALEKKVERIENNPRRRAKESKVREYYEK-----QFP-E 339
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              714 KMIAMGSAAALRNLMANR--PAKYKDANIMSPGSSLPSLHVRKQKALEAELDAQHLSETF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SSSSSRGSL----DSSRSEKDRSLERERGIG----LSTYHSATENPGTSSKRGLQLSAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             VGLLEYQHHSRDYASHLSPGSIIQPQRRRPSLLSEFQPGNERSQEL------HL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           615 VGTLTYRSQTNTLAIIESGGGIL ---RNVSSLIAT----NEDHRQILRENNCLQTLLQHL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                R--------PESHSYLPELGKSEM--EFIESKRPRLELLPDPLLRPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PLLATGOPAGSEDLTKDRSLTGKLEPVSPPSPPHTDPELEL-----VPPRLSKEELIONM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DRVDREITMVEQQISKLKKKQQQLEE---EAAKPPEPEKP-----VSP----P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DNID-NLSPKASHRSKQRHKQNLYGDYVFDASRHDDNRSDNFNTGNMTVLSPYLNTTVLP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               882 AAQIAKVMEEVSALHTSQDDRSPASAAELHCVAEERTAARRSSASHTHPNTHNFÄKSESS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 568 KSHSLTIVSNACGTLWNLSARNPKDQEALWDMGAVSMLKNLIHSKH-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            218 PIESKHRSLVQIIYDENRKKAEAAHRILEGLGPQVELPLYN---------
                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR008938; ARM.
InterPro; IPR000225; Armadillo.
Pfam; PF00514; Armadillo_seg; 4.
SMART; SM00185; ARM; 5.
PROSITE; PS50176; ARM REPEAT; 1.
Wht signaling pathway; Anti-oncogene; Phosphorylation; Coiled coil;
                        catenins. Binds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COLON TUMOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels 879;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 2842;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        C->R: IN AN IQ-INDUCED COLO
W; 3CBB2EA8A34E8F47 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ARM 1.
ARM 2.
ARM 3.
ARM 4.
ARM 5.
ARM 5.
ARM 7.
ARM 7.
ARM 7.
ARM 7.
ASP/GLU-RICH (ACIDIC).
ASP/GLU-RICH (ACIDIC).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL) (POTENTIAL)
                SUBUNIT: Forms homooligomers. Associates with axin (By similarity).
PTM: Phosphorylated by GSK3B (By similarity).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1; Score 355.5; DB 1; al Similarity 18.7%; Pred. No. 0.00016; 514; Conservative 334; Mismatches 1019;
phosphorylation state (By similarity)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HIGHLY CHARGED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COILED COIL
                                                                                                  SIMILARITY: Contains 7 ARM repeats.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LEU-RICH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          523
                                                                                                                                                                                                                                                                                                                                                              EMBL; D38629; BAA07609.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2831
1155
1575
1891
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             523 5;
2842 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ...........
                                                                                                                                                                                                                                                                                                                                                                                           Q02248; 3BCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1556
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       682
```

1833 RGGFT	Ω			1969	1535 KPRQSI	11: 2000 KPQAS	1587 2045 AMPKK	1611 -YEHLI		1659PTYE			1777 LSPGGE				1897 VLRSTS 2344 OLPRTS		1950 RFRADT ::: 2392 -PRSES						2557 SLPKVS			: 2648 APAVSR	2286 KKQEIN 2697 SKDTQG
a &	ු සි	δŏ	a &	∂ 8	λō	qa	<i>장</i> 옵	ð	qa	δ f	3 8	i 8	ò	අු	ò	g d	Š	3 8	중 요	ò	qq	<i>λ</i> ο	අු (ð t	g &	· 음	ò	qq	ර් සි
	400 YDADQQRIKFINMMGLMADPMKVYKDRQVMNNWSEQEKETFREKFMQHPKNFGLIASFLE 459 1059 KQNBQRQSRSQNTNPPVYSENTDKHLKFOOHFG 1092	KONENYKSLVRRSYRRGKSQQQQQQQQQQQOQPMPRSSQ	1093QQECVSPYRSRGTNGSETNRMGSSHAVNQNVNQSLCQEDDYEDDKFTNYS 1142	GEDNDEKEAVASKGR	11*3 brisbeechbeeerrinislataineekhyvöyldiskatalissSQKPSFSFS 1196 570 KTANSOGRRKGRITESMANBANSERATTPOOSARIASMEINESSPEMPEREMPERMPERMFAKKGIIE 600		DEILO	OHKLKMEKERNA	: : : : TRTKPSRLQASGLASESARH	716EMVEEAEALHASGNEVPRGECSGPATVNNSSDTESIPSPHTEAAKDTGQNGP 767	1347 KAVEFSSGÅKSPSKSGAQTPKSPPEHYVQETPLVFSRCTSVSSLDSFESRSIASSVÅSEP 1406	768 KPPATLGADGPPPGPPTPPRRTSRAPIEPTPASEATGAPTPPPAPFSPSAPPPVV 822	PKEEKEEETAAAPPVEEGEEQKPPAAEELAVDTGKAEEPVKSECTEEAEEGPAKGKDAE-		882AAEATAEGALKAEKKEGGSGRATTAKSSGAPQDSDSSATCSADEVDEAEGGDKNRL 937	1487 LLHFATESTPDG 1512	LSPRPSLLTPTGDPRAN	1513PFI	APPPPQNLQPESDAPOOPGSSPRGKSRSPAPPADKEAFAAE		1039 AQKLPGDPPCWTSGLPFPVPPREVIKASPHAPDPSAFSYAPFGHPLFLGLHDTARP 1094 1596 ASKLPPPVARKPSGLPVYKLLPSGSRLPSGAFYSFTPGND	VLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEHAKAPVGPVTMGL	1636 -VPRVYCVEGTPINFSTATSLSDLTIESPPNELAAGDGVRASVQSGEFEKRDTIPT 1690	1153 PLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEASVLRGTALGSVPGGS 1207	1691EGRSTDBAQRGKVSSIAIPDLDGSKAEEGDILAECINSALPRGR 1734			1262 SLPKGHVIYEGKKGHVISYEGGMSVTQCKKEDGRSSSGPPHETAAPKRIYDWMEGRVGRA 1321 1791 TDSKVNV	PRSYVEAQEDYL
qq	<i>≿</i> 8	ò	đ	ે દ	3 8	; A	કે ક	8 8	qq	ò	අ	දු ද	ò	đ	ờ	qq	ò	අ	δi	අ ,	8 8	ò	б	ò	අු	à	සු අ	දු දු	ò

183.	3 RGGFTFDSPHHYAPIEGTPYCFSRNDSLSSLDFDDDDV 1870
1372	2RREAKLLKREGTPPPPPBRDLTEAYKTQALGPLKLKPAHEGLVATVKEAGRSIHEI 1428
1429	9 PREELRHTPELPLAPRPLKE-GSITQGTPLKYDTGASTTGSKKHDVR 1474
1479	S SLIGSPGRTFPPVHPLDVMADARALERACYEESLKSRPGTASSSGGSIARGAPVIVPELG 1534
1535	5 KPRQSPLTYEDHGAPFAGHLPRGSPVTWREPTPRLQEGSLSS-SKASQDRKLT 1586
1587	7SSPHSTVPEHHPHPHSP 1610
1611	1 -YEHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLDAAAAYYLDRHLAPN 1658
1659	9PTYPHLYPPYLIRGYPDTAALENRQTIINDYITSQQMHHNTATAMAQRADMLRGLSPR 1716
1717	SESLALNYAAGPRGIIDLSQVPHLPVLVPPTPGTPATAMDRLAYLPTAPQPFSSRHSSSP 1776
1777	'LSPGGPTHLTKPTTTSSSERERDRDREREKSILTSTTTVEHAPIWRFGTEGSSG 1836 :
1837	SSGSSGGGGGSSSRPASHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGIITAVEPSKPT 1896
1897	
1950	RPRADTGHAFLAKPPARSGLEPASSPSKGSEPRPLVPPVSGHATIARTP 1998
1999	A-KNLAPHHASPDFPAPPASASDFHREKTOSKPFSIQELELRSLGYHGSSYSPBGVEPVS 2057 :
2058	PVSSPSLTHDKGLPKHLEELDKSHLEGELRPKQPGPVKLGGEAAHLPHLRPLPESQPS 2115
2116	SSPLLQTAPGVKGHQRVVTLAQHISEVITQDYTRHHPQQLSAPLPAPLYSFPGASCPVLD 2175
2176	LRRPPSDLYLPPPDHGAPARGSPHSEGGKRSPEPNKTSVLGGGEDGIEPV 2225
2226	SPPEGMTEPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQPPAFFSKLTESNSAMVKS 2285 :
2286	KKQEINKKLNTHNRNEPEYNISQPGTEIFNMPAITGTGLMTYRSQAVQEHASTNMGLEAI 2345

```
MIM; 114350;
                                                                                                                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                                                                                      DOMAIN
                                                                                                                                                                                                                                                           DOMAIN
8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    硆
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           . 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                        ------SKH--SSPSGTVAARV 2780
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DISEASE: Implicated in a subset of acute myeloid leukemia (acute myeloid controlymbrocytic leukemia) (AML) carrying a chromosomal translocation t(6;9) (p23;q34) that results in the formation of a DEK.CAN fusion gene.

DISEASE: Involved in some cases of acute undifferentiated leukemia (AUL) through a chromosomal translocation t(6;9) (q21;q34.1) that involves NUP214/CAN and SET.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DATABASE: NAME-Atlas Genet. Cytogenet. Oncol. Haematol.;
WWW="http://www.infobiogen.fr/services/chromcancer/Genes/CAN.html"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TISSUB-Testis,
MBDLNE-92195315; PubMed-1549122;
MBDLNE-92195315; PubMed-1549122;
Von Lindern M., Fornerod M., Van Baal S., Jaegle M., De Wit T.,
Bujis A., Grosveld G.;
"The translocation (6;9), associated with a specific subtype of acute myeloid leukemia, results in the fusion of two genes, dek and can, and the expression of a chimeric, leukemia-specific dek-can mRNA.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PTM: Probably glycosylated as it reacts with wheat germ agglutinin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      the cytoplasm.";

1. Vatl. Acad. Sci. U.S.A. 91:1519-1523 (1994).

1. FUNCTION: May serve as a docking site in the receptor-mediated import of substrates across the nuclear pore complex.

1. SUBUNIT: Homodimer. Interacts with DDX19 and NUPSB.

1. SUBCELLULAR LOCATION: Nuclear pore complex. Cytoplasmic filaments.

1. TISSUE SPECIFICITY: Expressed in thymus, spleen, bone marrow, kidney, brain and testis, but hardly in all other tissues or in whole embryos during development.
                        2346 IRKALMGKYDQWEESPPLSANAFNPLNASASLPAAMPITAADGRSDHTLTSPGGGGKAKV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE=94151361, PubMed=8108440;
Kraemer D., Wozniak R.W., Blobel G., Radu A.;
"The human CAN protein, a putative oncogene product associated with
myeloid leukemogenesis, is a nuclear pore complex protein that faces
                                                                                                                                                                                                                                                                                                                                      01-JUN-1994 (Rel. 29, Created)
01-JUN-1994 (Rel. 29, Last sequence update)
110-OCT-2003 (Rel. 42, Last annotation update)
Nuclear pore complex protein Nup214 (Nucleoporin) (CAN protein).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Buteleostomi,
Mammalia, Butheria, Primates, Catarrhini, Hominidae, Homo.
                                                                                                                           2437
                                                                                                                                                        ---PAP-GLASGDRPPSVSSVHSEG
                                                                                                                                                                                                                                                                                                           Z
                                                     Mol. Cell. Biol. 12:1687-1697(1992)
                                                                                                                           SG---RPSSRKAKS----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; X64228; CAA45535.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PIR; S26058; S26058.
Genew, HGNC:8064; NUP214.
                                                                                                                                                                                                                                                                                                           STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE PROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHARACTERIZATION,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    sapiens
                                                                                                                                                                                                                                                                                                           HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (WGA)
                                                                        2738
                                                                                                                           2406
                                                                     셤
                                                                                                                           ò
                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                           11D
DD411D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
D511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD51D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
DD511D
```

```
1098 -----RPPTISNPPPLIS-----SAKHPSVLERQIGA----ISQGMSVQ 1132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ----PPPQNLQPESD-APQQPGSSPRGKSR 1024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1025 SPAP-PADKEAFAAEAQKLP-----GDPP 1047
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1252 PSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDG----RSSSGPPHETAAP 1307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1308 -KRIYDMMEGRVGRAISSASIEGLMGRAIPPERHSPHHLKEQHHIRGSITQGI-PRSYVE 1365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1366 AQEDYLRREAKILKREGTPPPPPSRDLTEAYKTQALGPLK-----LKPAHEGLVATVKE 1420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1421 AGRSIHEIPREELRHTPELPLAPRPLKE---GSITQGTPLKYDTGASTTGSKKHDVRSLI 1477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1478 GSPGRTFPPVHPLD----VMADARALERACYEESLKSR--PGTASSSGGSIARGAPVIVPE 1532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----SQRYYEDL-----DEVSSTSSVSQSLESEDARTSCKDDEAVVQAPRHA 1016
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               192 ASVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDS 1251
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        808
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      608 PMSPFSSASKPAASGPLSHPTPLSAPPSSVPLKSSVLP---SPSGRSAQGSSSPVPSMVQ 664
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----IAHFQKELEELKARTSKACFQVGTSEEMKMLRTESDDLHTFLLE 753
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          973
                                                                                                                                                                                     Nuclear protein; Transport; Proto-oncogene; Chromosomal translocation; Repeat; Glycoprotein.

481 2076 11 X 5 AA APPROXIMATE REPEATS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1048 CWTSGLPFPVPPREVIKASPHAPDPSAFSYAPPGH----PLPLGLHDTARPVLP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LHVPYSEHAK-APVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SQAPQKLDASAAAAPASLPPSSPAAPIATFSLLPAGGAPTVFSFGSSSLKSSATVTGEPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   501 SYSSG-----ADSSKAAP-GPGPSTFSFVPPSKASLAPTP----AASPVAPSAASFS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 809 AQLQEIRRIHQYVKF----AVQDVNDVLDLEWDQHLEQKKKQRHLLVPERETLFNTLAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       918 NALLKTIESHTKSLPKVPAKLSPWKQAQLRNFLAKRKTPPVRSTAPASLSRSAFL---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       381 KTLPPAPVLMLLSTDGVLCPFYMINQNPGVKSLIKTPERLSLEGERQPKSPGSTPTTPTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        665 KSP-RITPPAAKPGSPQAKSL-QPAVAEKQGHQWKDS-----DPVMAG----IGEE-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -NREIINQORKRINHLVDSLQQLRLYKOTSLWSLSSAVP----SQSSIHSFDSDLESLC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   - IAKSPHSTVPEHHPHPISPY--EHLLR----GVSGVDLYRSH---IPLAFDPTSIPRGI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LGKPROSPLTYEDHGAPFAGHLPRGSPVTMREPTPRLOEGSLSSSKASQDRKLTSTPRE-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 2.7%; Score 355; DB 1; Length 2090; Best Local Similarity 20.6%; Pred. No. 0.00013; Matches 362; Conservative 177; Mismatches 656; Indels 560;
                                                                                                                                                                                                                                               11 X 5 AA APPROXIMATE REPEATS.
18 X 4 AA AA APPROXIMATE REPEATS.
11 X 3 AA APPROXIMATE REPEATS.
PRO/SER/THR-RICH.
LEUCINE-ZIPPER 1.
LEUCINE-ZIPPER 2.
BREAKPOINT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MW; 6DBE767FDD857F8F CRC64;
GO; GO:0005643; C:nuclear pore; TAS.
GO; GO:0005215; F:transporter activity; TAS.
InterPro; IPR004325; Nucleoporin_FG.
InterPro; IPR001680; WD40.
                                                                                                                         Pfam; PF03093; Nucleoporin_FG; 22
SMART; SM00320; WD40; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        213766
                                                                                                                                                                                                                                                                                  2084
                                                                                                                                                                                                                                                                                                               2085
                                                                                                                                                                                                                                                                                                                                                                                                     861 88
812 81
2090 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          993 KPAPPAP--
                                                                                                                                                                                                                                                                                  1409
1427
1213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                441
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  710
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
```

```
Gallus
                                                                                                                                                                                                                                                                                                                                                                                                                                  Ξ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9
DDTTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDD
                                                                                                                                                       1775
                                                                                                                                                                                                                                                            1250
                                                                                                                                                                                                                                                                                            GSSGSSGGGGSSSRPASHSHAHQHSPISPRIQDALQQRPSVLHNTGMKGIITA--VEPS 1893
                                                                                                                                                                                                                                                                                                                             1251 QPDAFSSGGG---SKP-SYEAIPESSP------PS-----GITSASNTTPG 1286
                                                                                                                                                                                                                                                                                                                                                                                                EP----AASSSRPVAPSGTALSTTSSKL----ETPPSKLGELLPPSSLAGETLGSFSGL 1337
                                                                                                                                                                                                                                                                                                                                                                                                                                    ---EAPRVARPERPRADT------GHAFLAKPPARSGLEP--ASSPSKGSE 1980
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RVGOADDSTKPTNKASSTSLTSTQPTKTSGVPSGFNFTAPPVLGKHTEPPVTSSATTTSV 1397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ----APHHASPDPPA--PP 2015
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1398 APPAATSTSSTAVFGSLPVTSAGSSGVISFGGTSLSAGKTSFSFGSQQTNSTVPPSAPPP 1457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2016 ASASDPHREKTQSKPFSIQELELRSLGYHGSSYSPEGVEPVSPVSPSPSLTHDKGLPKHLE 2075
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ELDKSHLEGELRPKQPGPVKLGGEAAHL---PHLRPLPESQPSSSP-----LLQTAPG 2125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EATSSAL----PEKPGDSEVSASAASLLEEQQSAQLPQAPPQTSDSVKKEPVLAQPAVS 1549
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1550 NSGTAASSTSLVALSAEATPATTGVPDARTEAVPPASSFSVPGQTAVTAAAISSAGPVAV 1609
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1610 ETSSTPIAS---STTSIVAPGPSAEAAAFGTVTSGSSVFAQPPAASSSSAFNQLTNNTAT 1666
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PPEGMTEPGHSRSAVYPLLYRDGEQT------EPSRMGSKSPGNTSQPPAFFSKL 2275
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         T-----TEISNSAMVKSKKQEINKKLNTHNRNEPEYNISQPG---TEIFNMPAITGT 2322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----SLPAAMPITAADGRSDHTLTSPGGGGKAKVSGRPSSRKAKSPAPGLASGDRP 2427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1813 GQGGSVFGGTSAATTTAATSGFSFCQASGFGSSNTGSVFGQAAS-----TGGTVFGQQS 1866
                  -TIIN 1687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1667 APSATPVFGQVAASTAPSLF--GQQTGSTASTAAATPQVSSSGFSSPAFGTTAPGVFGQT 1724
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TFGQASVFGQSASSAASV----S 1761
                                    --DYITSQOMHHNTATAMAQRADMLRGLSPRESSLALNYAAGPRGIIDLSOVPHLPVLVP
                                                                                                          --PTPGTPATAM------DRLAYLPTAPQPFSSRHSSS
                                                                                                                                                                                       1158 VAANQAKQGSLINSLKPSGPTPASGQLSSGDKASGTAKIETAVTSTPSASGQFSKPFSFS
                                                                                                                                                                                                                          PLSPGGPTHLTXPTTTSSSERERDRDRERDRDREREKSILTSTTTVEHAPIWRPGTEQSS
                                                                                                                                                                                                                                                            -----NFTAAQGA---TPSTKESS
                                                                                                                                                                                                                                                                                                                                                             KPTVLRSTSTSSPVRPAATFPPATHCPLGGTLDGVYPTLMEPVLLPK------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STPTSTSGSVFGAASST-----SSSSSFSFGQSSPNTGGGLFGQSNAPAFGQSPGF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                --ASCPVLDLRRPPSDLYLPPPDHGAPARGSPHSEGGKRSPEPNKTSVLGGGEDGIEPVS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2323 GLMTYRSQAVQEHASTNMGLEAIIRKALMGKYDQWEESPPLSANAFNPLNASA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2126 VKGHQRVVTLAQHISEVITQDYTRHHPQQLSAPLPAPLYSFPG-------
                PLDAAAAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAALENRQ-
                                                                                                                                                                                                                                                        PSGTGFNFGIITPTPSS------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1981 PRPLVPPVSGHATIARTPAKNL---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1881
                                                                                                                 1119 VPQVVNVQELKNNPATP--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2428 PSVS-SVHSEGDCNR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SSSSGSVFGSGNTGR
                                                                                                                                                                                                                                                                                                                                                               1894
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1338
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1762
              1642
                                                                                  1688
                                                                                                                                                     1746
                                                                                                                                                                                                                          1776
                                                                                                                                                                                                                                                          1218
                                                                                                                                                                                                                                                                                                                                                                                                 1287
                                                                                                                                                                                                                                                                                                                                                                                                                                    1941
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2076
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1725
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2376
                                                101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2276
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1867
                                                                                                                                                                                                                                                                                              1836
                                                                                 ò
                                                                                                          셤
                                                                                                                                            ò
                                                                                                                                                                                    음
                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ठ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      용
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
```

1453 AA.

PRT;

STANDARD;

RESULT 50 CA11_CHICK ID _CA11_CHICK

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the FWBL outstation the Buropen Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     end
                                                                                                                                                                                                     oailus galus (ultoren).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  carboxyterminus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 cken pro-
acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          hydroxyapatite.

PTM: Proline residues at the third position of the tripeptide repeating unit (G-X-Y) are hydroxylated in some or all of the chains. Pro-1153 is the only 3-hydroxypro and the only hydroxylated proline in position X. SIMILARITY: Contains 1 WWFC domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ú
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE=80134546; Pubperson.

MEDLINE=80134546; Pubperson.

Showalter A.M., Pesciotta D.M., Eikenberry E.F., Yamamoto T.,

Pastan I. Decrombrugghe B., Fietzek P.P., Olsen B.R.;

"Nucleotide sequence of a collagen cDNA-fragment coding for the
carboxyl end of pro alpha 1(I)-chains.";

FES Lett. 111:61-65(1980).

-I- FUNCTION: Type I collagen is a member of group I collagen

(fibrillar forming collagen).

-I- SUBUNT: Trimers of one alpha 2(I) and two alpha 1(I) chains.

-I- SIBSUE SPECIFICITY: Forms the fibrils of tendon, ligaments an
bones. In bones the fibrils are mineralized with calcium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE OF 981-1453 FROM N.A.
MEDLINE=81160715; PubMed=6927845;
Fuller F., Boedtker H.;
"Sequence determination and analysis of the 3' region of chicken alpha 1(I) and pro-alpha 2(I) collagen messenger ribonucleic acid including the carboxy-terminal propeptide sequences.";
Blochemistry 20:996-1006(1981).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE OF 152-1187.
MEDLINES-82231995; PubMed=7093229;
Highberger J.H., Corbett C., Dixit S.N., Yu W., Seyer J.M.,
Kang A.H., Gross J.;
"Amino acid sequence of thick skin collagen alpha 1(I)-CBB and
complete primary structure of the helical portion of the chick
collagen alpha 1(I) chain.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 clones encoding the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE OF 1-144 FROM N.A.

MEDLINE=88007542; PubMed=2820966;

Finer M.H., Aho S., Gerstenfelld L.C., Boedtker H., Doty P.;

"Unusual DNA sequences located within the promoter region and first intron of the chicken pro-alpha 1(1) collagen gene.";

J. Biol. Chem. 262:13323-13332(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  at
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE=72243016; PubMed=5047697;

Eyre D.R., Glimcher M.J.;
"Evidence for a previously undetected sequence at of the alpha 1 chain of chicken bone collagen.";
Biochem. Biophys. Res. Commun. 48:720-726(1972).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            of the chicken pro alpha 1(I) collagen mRNA.";
Gene 56:71-78(1987).
                                     01-OCT-1989 (Rel. 12, Last sequence update)
10-OCT-2003 (Rel. 42, Last annotation update)
Collagen alpha 1(I) chain precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CDNA
                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE-88056316; PubMed-3678834;
Finer M.H., Boedtker H., Doty P.;
"Construction and characterization of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Biochemistry 21:2048-2055(1982).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE OF 1311-1453 FROM N.A.
   (Rel. 01, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         [5]
SEQUENCE OF 981-1453 FROM N.
                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE OF 1-153 FROM N.A.
                                                                                                                                                                              Gallus gallus (Chicken)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE OF 1200-1205
                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=9031;
21-JUL-1986
01-OCT-1989
```

```
1408 KPAHEGLVATVKEAG----RSIHEIPREELRHTPELPLAPR-----PLKEGSITQGTPLK 1458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1459 YDTGASTTGSKKHDVRSLIGSPGRTFPPVHPLDVMADARALERACYEESLKSRPGTASSS 1518
                                                                                                                                                                                                                                                                                                                                                                    1359 IPRSYVEAQEDYLRREAKLLKREGTPPPPPPSRDLTEAYKTQALGPLKL------ 1407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1519 GGSIARGAPVIVPELGKPROSPLTYEDHGAPFAGHLPRGSPVTMREPTPRLOEGSLSSSK 1578
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1579 ASQDRKLTSTPREIAKSPHSTVPEHHPHPISPYEHLLRGVSGVDLYRSHIPLAFDPTSIP 1638
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2041 LGYHGSSYSPEGVEPVSPVSPSTTHDKGLPKHLEELDKSHLEGELRPK----QPGPVKL 2096
                                                                                                                                                                                                                                                                                                 1309 RIYDMMEGRVGRAISSASIEGLMG-----RAIPPERHSPHHLKEQHHIRGSITQG 1358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1639 RGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAALENRQTIINDYITSQQMHHN 1698
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1759 AYLPTAP-QP-----FSSRHSSSPLSPGGPTHLTKPTTTSSERERDRDRERDRDRE 1809
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1810 REKSILTSTTTVEHAPIWRPGTEQSSGSSGSSGGG--GSSSRPASHAHQHSPISPRT 1867
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1868 QDALQQRPSVLHNTGMKG----IITAVEPSKPTVLRSTSTSSPVRPAATFPPATHCPLGG 1923
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1924 TLDGVYPTLMEPVLLPKEAPRVARPERPRADTGHAFLAKPPARSGLE-PASSPSKGSEPR 1982
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----KNGDRGETGPAGPAGPPGPAGA 1072
                                                                                                                                                                                                                            1252 PSRLDRGREDSLP--KGHVIYEGKKGHVLSYEGGMSVTQCSKEDGR-SSSGPPHETAAPK 1308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1132 QLHVPYSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TATAMAQRADMLRGLSPRESSLALNYAAGPRGIIDLSQVPHLPVLVPPTPGTPATAMDRL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             --LHDTARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSV
                                                                                                                       ------GPAGERGAPGSRGFPGAD-
                                                                                                                                                                                                                                                                                                                                  -----GARGQA-----GVMGFPGPKGAAGEPGKPGERGAP------G
                                                                                                                                                                                                                                                                                                                                                                                                     592 PPGAVGAÁGKD---GEÁ---GAQGPPGPTGPAGERGEQGPAGAPGPGGPAGPPGEAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                               ---DRG-----DPGPKGADG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -----GZVGAPGPKGARGSAGPPGATGFPGAAGRV------GPPGPSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NIGLPGPPGPAGKZGSKGPRGETGPAGRPGEP-----GPAGPPGEFKGSPGADGPIGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PLVPPVSGHATIARTPAKNLAPHHASPDPPAPPA--SASDPHREKTOSKPFSIQELELRS
                                                                                                                                                         1192 ASVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDS
                                                                                                                                                                                          -----GIAGPKGPPGER------GSPGAVGPKG-----S
                                                                                                                                                                                                                                                             512 PGBAGRPGBAGLPGAKGLTGSPGP----DGKTGPPGPAGQDGRPGPAGPP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                699 GDAGAP-----GAPGNEGPP-----GAPGNEGPAA
                                                    -----GPAGEEGKRG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PMGPP----GLAGPPGE--AGREGAPGAEGAPGRDGAGPKGDRGETGP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AGPPGADGÓPGAKGETGDAGAKGDAGPPGPAGPTGAPGPA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         764 AGAPGDKGEAG-----PPGPAGP-----TGA-
                                                    415 APGPKGNSGEPGAPGNKGDTGAKGEPGPAGVQGPP-
                                                                                                                       -----ARGEPGPA--GLP-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AGPPGAPGAPGAPGPVGPAG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GLPGAKG
                   1082 HPLPLG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1983
                                                                                                                                                                                            489
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      745
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 803
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     843
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              886 ...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1031
                                                                                                                                                                                      유
                                                                                                                                                                                                                                                    쉱
                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                 ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        유
                                                                                                                셤
                                                                                                                                                      8
                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              윤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                           셤
                                                                                     ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    . <u>a</u>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QVTKVHEPPREDAAPTKPA-PPAPPPQNLQPESDAPQQPGS-SPRGKSRSPAPP---- 1029
   commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           85;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----ADKEAFAAEAQKLPGDPPCWTSGLPFPVPPREVIKASPHAPDPSAFSYAPPG 1081
                  (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                813
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GPRGDRGLPGPPGRDGIPGQPGLPGPPGPPGLGGNFAPQMSYGYDEKSAGVAVPGPM 170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GPAGPRGLPGPPGAPGPQGFQ----GPPGEPGE----PGASGPMGPRGPAGPPGKN---- 218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSSGAPQDSDSSATCSADEVDE 928
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AEG-----GDKNRLLSP----RPSLLTPTGDP--RANASPQKPLDLK-QLKQRAAAIPPI 976
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AKGOPGPAGPKGEPGSPGENGAPGQMGPRGLPGERGRPGPSGPAGARGNDGAPGAPGP- 323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ------GPTGPAGPPGFPGAAGAKGETGPQGARGSEGPQGSRGEPGPPGPAGA 370
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -----GDDGEAGKPGRPGQRGPPGPQGARGLPGTAGLPGMKGHRGFSG-----LDG 264
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     371 AGPAGNPGADGQPGAKGATGAPG----IAGAP------GFPGARGPSG-PQGPSG 414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SGN----EEEMVEEAEALHASGNEVPRGECSGPATVNNSSDTESIP----SPHTEAAK-DT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GONG----PKPPATLGADGP----PPGPPTPPRRTSRAPIEPTPASEATGAPTPPPAPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   814 SPSAP----PPVVPKEEKEEETAAAPPVEEGEEGKPPAAEELAVDIGKAEEPVKSECTE
                                                                                                                                                                  InterPro; IPR008161; Clg helix.

BR InterPro; IPR008161; Collagen.

BR InterPro; IPR008161; Collagen.

BR InterPro; IPR001818; Fibrinogen.C.

BR InterPro; IPR001019; WWP.C.

BR InterPro; IPR001019; WWP.C.

BR Fam; PF01391; Collagen; 18.

BR Pf01391; Collagen; 18.

BR Pf01391; Collagen; 18.

BR Pf01391; Collagen; 18.

BR Pf01391; Collagen; 18.

BR Pf01391; Collagen; 18.

BR Pf01391; Collagen; 18.

BR Pf01391; Collagen; 18.

BR SMART; SM00214; WWC; 1.

BR ROSITE; PS01208; WWC; 1; 1.

BR ROSITE; PS0184; WWC; 2; 1.

BR Extracellular matrix; Connective tissue; Repeat; Hydroxylation; W Glycoprotein; Collagen; Signal; Pyrrolidone carboxylic acid.

TR SGNAL
   for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ilarity 20.8%; Pred. No. 0.0001;
Conservative 110; Mismatches 584; Indels 678;
   and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2.7%; Score 353; DB 1; Length 1453;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PYRROLIDONE CARBOXYLIC ACID.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -> L (IN REF. 5).
-> H (IN REF. 6).
3BC6152134271F4D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HYDROXYLATION (POTENTIAL).
HYDROXYLATION (POTENTIAL).
HYDROXYLATION (POTENTIAL).
HYDROXYLATION.
F -> L (IN REF. 5).
Q -> H (IN REF. 6).
   Usage by
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AMINO-TERMINAL PROPEPTIDE.
COLLAGEN ALPHA 1(1) CHAIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COLLAGEN ALPHA 1 (I) CHAIN C-TERMINAL PROPEPTIDE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HYDROXYLATION (POTENTIAL)
modified and this statement is not removed.
                   entities requires a license agreement (
or send an email to license@isb-sib.ch)
                                                                                                              EMBL; MI0571; AAA48671.1; ALT.SEQ.
EMBL; MI7607; AAA48672.1; -..
PIR; A27179; A27179.
PIR; I50629; I50629.
                                                                 EMBL; M17839; AAA48704.1; -.
EMBL; M17838; AAA48704.1; JOINED.
EMBL; V00401; CAA23695.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1153 1153 HYDI
1187 1187 F -:
1441 1441 Q -:
1453 AA; 137789 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          151
1205
1453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches 360;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        869
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MOD_RES
CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            711
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                763
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           929
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         219
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                977
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
PROPEP
DOMAIN
MOD RES
MOD RES
MOD RES
MOD RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local
8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ठे
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8 8 8 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ઠે
```

645

us-09-522-753-5.rsp

```
2097 GGEAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQRVVTLAQHISEVITQDYTRHHPQQLS 2156
                                                                                    1101 -----GFSG-----LQGPPG----PP--GAPGEQGPSGASGPAGPRGPPGSAGA 1138
                                                                                                                                           1139 AĞKDĞLNGLPGPIĞPPGPRGRTGEVGPV----ĞPPGPPGPPGPPGPPGFPSGGFDFSFLPQPP 1194
                                                                                                                                                                                                                                                                                    O9EPQ8; Q60792;
28-FEB-2003 (Rel. 41, Created)
28-FEB-2003 (Rel. 41, Last annotation update)
28-FEB-2003 (Rel. 41, Last annotation update)
Transcription factor 20 (Stromelysin 1 PDGF-responsive element-binding
protein) (SPBE-binding protein) (Nuclear factor SPBP).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  INTERACTION WITH RNF4, TISSUE SPECIFICITY, AND MUTAGENESIS.
MEDLINE=20408957; PubMed=10849425;
Lyngsoe C., Bouteiller G., Dangaard C.K., Ryom D., Sanchez-Munoz S., Noerby P.L., Bonven B.J., Oeergensen P.;
"Interaction between the transcription factor SPBP and the positive cofactor RNF4. An interplay between protein binding zinc fingers.";
J. Blol. Chem. 275:2644-26149(2000).
-! FUNCTION: Transcriptional activator that binds to the regulatory region of MMP3 and thereby controls strongelysin expression. It stimulates the activity of various transcriptional activators such as JUN, SPI, PAX6 and ETS1, suggesting a function as a
                                                         2157 APLPAPLYSFPGASCPVLDLRRPPSDLYLPPPDHGAPARGSPHSEGGKRSPEPNKTSVLG
                                                                                                                2217 GGEDGIEPVSPPEGMTEPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGN-----TSQPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SUBDUIT: Homodimer (Probable). Interacts with RNF4 and JUN. Binds to the regulatory region of MMP3. SUBCELLULAR LOCATION: Nuclear.
ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN=BALB/C; TISSUE=Brain; MEDLINE=20568288; PubMed=10995766; Rekdal C., Sjoettem E., Johansen T.; The nuclear factor SPBP contains different functional domains and stimulates the activity of various transcriptional activators."; J. Biol. Chem. 275:40288-40300(2000).
                                                                                                                                                                       2270 -----AFFSKLTESNSAMVKSKKQEINKKLNT-----HNRNEPEYNISQP 2309
                                                                                                                                                                                         Mus musculus (Mouse),
Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia, Eutheria, Rodentia, Sciurognathi, Muridae, Murinae, Mus.
NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDILINE-56279378; PubMed-8663478;
Kirstein M., Sanz L., Moscat J., Diaz-Meco M.T., Saus J.;
Kirstein M., Sanz L., Moscat J., Diaz-Meco M.T., Saus J.;
"Cross-talk between different enhancer elements during mitogenic induction of the human stromelysin-1 gene.";
J. Biol. Chem. 271:18231-18236(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TISSUE=Fibroblast;
WBDINDS=95280915; PubMed=7760812;
Sanz L., Moscat J., Diaz-Meco M.T.;
"Molecular characterization of a novel transcription factor that
                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), CHARACTERIZATION, AND
                            1073 RGPAG-----PQGPRGDKGETGEQGDRGMKGHR-
                                                                                                                                                                                                                                                                           PRT; 1983 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE OF 774-1965 FROM N.A. (ISOFORM 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            controls stromelysin expression."; Mol. Cell. Biol. 15:3164-3170(1995).
                                                                                                                                                                                                                                                                           STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      INTERACTION WITH JUN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ALTERNATIVE SPLICING.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 coactivator.
                                                                                                                                                                                                                                                                           MOUSE
                            요
                                                                                 ద
                                                                                                             ò
                                                                                                                                        a
                                                                                                                                                                      ò
                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                       ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              91;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A->T: LOSS OF INTERACTION WITH RNF4; WHEN ASSOCIATED WITH S-1702, R-1736 AND V-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 P->S: LOSS OF INTERACTION WITH RNF4; WHEN ASSOCIATED WITH T-1629, R-1736 AND V-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CG-SKV: LOSS OF INTERACTION WITH RNF4;
WHEN ASSOCIATED WITH T-1629 AND S-1702.
C->A: REDUCES THE INHIBITORY EFFECT OF
THE ATYPICAL PHD DOMAIN.
C->A: REDUCES THE INHIBITORY EFFECT OF
THE ATYPICAL PHD DOMAIN.
THE ATYPICAL PHD DOMAIN.
C->A: REDUCES THE INHIBITORY EFFECT OF
THE ATYPICAL PHD DOMAIN.
C->A: REDUCES THE INHIBITORY EFFECT OF
THE ATYPICAL PHD DOMAIN.
THE ATYPICAL PHD DOMAIN.
THE ATYPICAL PHD DOMAIN.
                                                                                                                                                                                                  -I- DOMAIN: The arypical PHD domain functions as a negative modulator of cofactor binding.
-I- SIMILARITY: Contains 1 A.T hook DNA-binding repeat.
-I- SIMILARITY: Contains 1 PHD-type zinc finger.
-I- CAUTION: Ref.2 sequence differs from that shown due to
                                                                                                                                 DEVELOPMENTAL STAGE: Isoform 2 is exclusively expressed at 7-11 days of development. Isoform 1 is found only at low levels in 15-17 days embryos.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                        lung, liver, kidney and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; AX007594; AAG28929.1; -.
EMBL; U20282; AAA86495.1; ALT_SEQ.
MGD; MGI:108399; Tcf20.
GO; GO:0005634; C:nucleus; ISS.
GO; GO:0003677; F:DNA binding; ISS.
GO; GO:0003713; F:transcription co-activator activity; ISS.
GO; GO:0006355; P:regulation of transcription, DNA-dependent; ISS.
InterPro; IPR001965; Znf_PHD.
SMART; SM00249; PHD; I.
Transcription regulation; Activator; DNA-binding; Zinc-finger;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CPLPPLQNKTAKGSLSTEQSERG -> VRLWR (in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2.7%; Score 353; DB 1; Length 1983; ilarity 18.7%; Pred. No. 0.00014; Conservative 240; Mismatches 731; Indels 732;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LEUCINE-ZIPPER.
NUCLEAR LOCALIZATION SIGNAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A.T HOOK.
NUCLEAR LOCALIZATION SIGNAL.
NUCLEAR LOCALIZATION SIGNAL.
PHD-TYPE (ATYPICAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          003986.
                                        Name=2;
IsoId=Q9EPQ8-2; Sequence=VSP_003986;
TISSUE SPECIFICITY: Expressed in brain,
Name=1;
IsoId=Q9EPQ8-1; Sequence=Displayed;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SER-RICH.
POLY-GLN.
POLY-PRO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Nuclear protein, Alternative splicing.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POLY-ALA.
POLY-GLN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POLY-GLU.
POLY-GLU.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1983 AA; 215284 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1295
1579
1628
1819
1962
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1629
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1702
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1737
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1926
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1936
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1939
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1931
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity
Matches 392; Conserv
                                                                                                                                                                                                                                                                                                                       frameshifts.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1604
1812
1911
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1939
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1681
1793
1961
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1629
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1926
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1936
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1702
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1931
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1736
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DNA BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MUTAGEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MUTAGEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MUTAGEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MUTAGEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MUTAGEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MUTAGEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MUTAGEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
```

RESULT 52 CA15_HUMAN CA15_HUMAN AC P20908; DT 01-FEB-1991 (Rel. 17, Created)	CA15	1331 MGRA	\$ 8 8
2210NKTSVLGGGBDGIBPVSPPEGMTEPGH ::	ờ ፭ ፡ ;	1291	& 43
2169 ASCPVLDLRRPPSDLYLPPPDHGAPARGSP	γο dg	EDSPSRLDRGREDS - LPKGHVIYEGKKGHVLSYEGG MSVTQCS MSVTQCS MSVTQCS	8 &
2114 PSSSPILQTAPGYKGHQRVVTLAQHISBVITQ -	රු අ <u>අ</u>	1191 EASVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYKGTITRIIG 1248	& 43
KTDAKNKSFFPYIHVVNKCELGAVCT	S 6	1136 PYSEHAKAPVGPVTMGLPLPMDPKKTAPFSGVKOEOLSPRG-QAGPPESLGVPTAQ 1190 	& A
1600 DGEPKPKKQRQRRERRKPGAQPRKRKTKQAV 2079 KSHLEGRCALTRPKOPGPV	a 8		e d
SDPHREKTQSKPFSIQELELRSLGY	8 8		q
1963 PPARSGLEPASSPSKGSEPRPLVPPVSGHATIARTE	<i>à</i> €	KBAFAABAQKLPGDPPCWTSGLPPPVPPRBVIKASPHAPDPSAF	ò
1907 VRPAAIFPPAIHCPLGGTLDGVYPTLMEPVLL 	충 옵 	994 PAPPAPPPPQNLQPESDAPQQPGSSPRGKSRSPAPPAD 1031	કે ક
	ପ ୍ର	934 KNKLLSPRESLLITGJPRAMASPQKRELDKQLKQRAAALPPIQVIKVHEPPREDAAPTK 993 	<u>8</u>
1429	8 &		a
1794 SERERDRDREREREKSILTSTTVEHAPI	i i		B
1/34 DSQVFTDFVLVEFIFGIFFIFTHATUREALFIFTHATQFFF 	ਤੇ <u>ਜ</u> ਼	820 PVVPKEEKEEETAAAPPVEEGEEQKPPAABELAVDTGKAEEPVKSECTEEAEEGPAKGKD 879	કે ક
	ପ୍ର	782PPTPPRRTSRAPIEPTPASEATGAPTPPAPPSPSAPP 819	å 5
1331SHSODIKSIPKRDSSKDLPNPDNR	셤 . 6	: 428 QGNVPMSSRNRILQLLPQLSPTPSMMPSPNSHAAGFKGFGLEGVPEKRLTDPGLSSLSAL 487	q
1617 GVSGVDLYRSHIPLAFDPTSIPRGIPLDAAAAX : : 	λ ₀	: :	3 8
1564EPIP KLQEGSJSKAQŪKKITSTPKEI 1284 RVRSFISPIPSKRQSQDVKNSNADDKGRLLHPSKEC	Š 8	709GVSGNEEBAVEEALHASGNEVPRGECSGPATVNNSSDTESIPSPHTEAAKDTG 763	हें है
QKRYG	ସ୍ପ ଓ	:	2 8
1518SGGSIARGAPVIVPELGKPRQSF	ò	209 GUSTUGSTSVNAGSUTEGHNVGSNAQAYGTQSNTSYQFQSM 309 660 ONLDEILOOHKLKWEKERNARRKKKKAPAAASEBAAPPPVVEDEBNEAS 708	8 8
1464 SIIGSANIUVROLIGSGANIFEVNIELDVANDARAL 1180 YHDPSTQEAGRCLMSSDGLPAQSMELKHSSQKU	· 음	ETAKKGLLEHGRNWSAIARMVGSKTVSQCKNFYFNYKKR	8 8
	셤 (ପ୍ର
1410AHEGLVATVKEAGRSIHEIPREELRHTPELPI	ò	166 YQQQASSQQQQQQQQQQQQQQQQQQVQQLRQQLYQSHQPLP 208 547 XEXTDDTSGEDNDEKEAVASKGRXTANSOGRPKGRTTRSMANRANSERITPOOSAEL 604	8 8
1079PFSPNSESLASAYHTNTRAHAYGDE	କ୍ଷ 	54	8

LLPKEAPRVARPERPRADTGHAFLAK 1962 TKLVRSRKGQRSLTPPPSSTESKVL 1710 SPHSEGGKRSPEP----- 2209 DUTGLNSQLHYKRQMYQQQQEEYKD 1127 PLAPRPLK -- EGSITOGTPLKYDTGA 1463 DRVRSPLKNDKDGMMYGPPV----GT 1179 ALERACYEESLKSRPGTASS---- 1517 |: :: | :: | :: | KLQESRWDLSRQTSPAKSSGPPGMSN 1236 SPLTYEDHGAPFAGHLPRGSPVTMR- 1563 ||| : :|| ----EDHSSQ-----NPLIMRR 1283 EIAKSPHSTVPEHHPHPISPYEHLLR 1616 AYYLPRHLAPNPTYPHLYPPYLIRGY 1673 DMLRGLSPRESSLALNYAAGPRGIID 1733 FSSRHSSSPLSPGGPTHLTKPTTTSS 1793 PIWRPGTEQSSGSSGSGGGSSSR 1850 |: || : | PL--PGPSBEWRGSG-----DDK 1472 SIITAVEPSKPTVLRSTSTSSP---- 1906 ---TGKEPS--GTMTSTASQKPGGNQ 1504 | |: |: |: LAPEANPKAEEKEN---DT---VMIS 1550 FPAKNLAPHHASPDPPAPP----ASA 2018 SKKQQQQP----PPPPQPPQMPEGSA 1599 VEPVSPVSSPSLTHDKGLPKHLEELD 2078 FOD-YTRHHPQQLSAPLPAPLYSFPG 2168 KSASNGSKTDTEEEEQQQQKEQRSL 1807 PVKLGGEAAHLPHLRPLPESQ---- 2113 SHSRSAVYPLLYRDGEQTEPS 2255 SSEKTV-----SDTKPS 1854

```
MEDLINE=90366601; PubMed=2203476;
Yaol Y., Hashimoto K., Koltebashi H., Takahara K., Ito M., Kato I.;
"Primary structure of the heparin-binding site of type V collagen.";
Biochim. Biophys. Acta 1035:139-145(1990).
                                                                                                                                                          Kato I.;
"Complete primary structure of human collagen alpha 1 (V) chain.";
J. Biol. Chem. 266:13124-13129(1991).
                                                 Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                      Seyer J.M., Kang A.H.;
"Covalent structure of collagen: amino acid sequence of three
cyanogen bromide-derived peptides from human alpha 1(V) collagen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE OF 565-576; 756-772; 1012-1029; 1219-1232 AND 1465-1477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TISSUE=Chorioamniotic membrane;
MEDLINE=94237164; PubMed=8181482;
Moradi-Ameli M., Rousseau J.C., Kleman J.P., Champliaud M.F.,
Boutillon M.M., Bernillon J., Wallach J.M., van der Rest M.;
"Diversity in the processing events at the N-terminus of type-V
collagen.";
                                                                                                                   SEQUENCE FROM N.A., AND SEQUENCE OF 556-565.
MEDLINE=91302336; PubMed=2071595;
Takahara K., Seto Y., Okasawa K., Okamoto N., Noda A., Yaoi Y.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     "Mutations in the COLSA1 gene are causal in the Ehlers-Danlos syndromes I and II.", Am. J. Hum. Genet. 60:547-554(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mann K.;
"Isolation of the alpha 3-chain of human type V collagen and characterization by partial sequencing.";
Biol. Chem. Hoppe-Seyler 373:69-75(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DISEASE, AND VARIANT EDS-I SER-1639.
MEDLINE=97195540; PubMed=9042913;
de Paepe A., Nuytinck L., Hausser I., Anton-Lamprecht I.,
Nacygert J.-M.;
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                             Arch. Biochem. Biophys. 271:120-129(1989)
                                                                                                                                                                                                                                                                                                                                      SEQUENCE OF 823-950, AND HEPARIN-BINDING.
                         Collagen alpha 1(V) chain precursor.
                                                                                                                                                                                                                              TISSUE=Chorioamniotic membrane;
MEDLINE=89227189; PubMed=2496661;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Eur. J. Biochem. 221:987-995(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE=92239022; PubMed=1571108;
                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE OF 556-571.
                                                                                                                                                                                                                 SEQUENCE OF 621-822
                                                                                            NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                   TISSUE=Placenta;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               collagen.
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation. The European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
PTM: Sulfated on 40% of tyrosines.

DISBARS: Defects in COL5A1 are a cause of Ehlers-Danlos syndrome type I (EDS-I) [MIM:130000]; also known as Ehlers-Danlos syndrome gravis. EDS-I is a connective-tissue disorder characterized by loose-jointedness and fragile, velvety, stretchable, bruisable skin that heals with peculiar 'cigarette-paper' scars. Inheritance is autosomal dominant.
                                                                                                                                                 -!- DISBASE: Defects in COLSAl are a cause of Bhlers-Danlos syndrome type II (BDS-II) [MIM:13010]; also known as Bhlers-Danlos syndrome mitis. Inheritance is autosomal dominant.
-!-SIMILARITY: HIGH, TO ALPHA 3(V) AND ALPHA 1(XI) CHAINS.
-!-SIMILARITY: CONtains 1 TSP N-terminal (TSPN) domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SMART; SM00282; LamG; 1.
SMART; SM00210; TSPN; 1.
Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Collagen; Signal; Heparin-binding; Sulfation; Disease mutation;
Ehlers-Danlos syndrome; 3D-structure.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NONHELICAL REGION.
CARBOXYL-TERMINAL PROPEPTIDE.
SULFATION (POTENTIAL).
SULFATION (POTENTIAL).
SULFATION (POTENTIAL).
SULFATION (POTENTIAL).
SULFATION (POTENTIAL).
SULFATION (POTENTIAL).
SULFATION (POTENTIAL).
SULFATION (POTENTIAL).
SULFATION (POTENTIAL).
SULFATION (POTENTIAL).
SULFATION (POTENTIAL).
SULFATION (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NONHELICAL REGION.
INTERRUPTED COLLAGENOUS REGION
TRIPLE-HELICAL REGION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COLLAGEN ALPHA 1 (V) CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TSP N-TERMINAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HYDROXYLATION.
HYDROXYLATION.
HYDROXYLATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HYDROXYLATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HYDROXYLATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HYDROXYLATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MIM; 130010; -
GO; 6000558; C:collagen type V; TAS.
INCEPPO; PR008161; Clg helix.
InterPro; IPR008160; Collagen.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SULFATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR008985; ConA_like_lec gl.
InterPro; IPR000885; Fib collagen_C.
InterPro; IPR001791; Laminin_G.
InterPro; IPR003129; TSPN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ProDom; PD000007; Cig_helix; 1.
ProDom; PD002078; Fib_collagen_C; 1.
SMART; SM00038; COLFI; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; D90279; BAA14323.1; -. PDB; 1A89; 18-NOV-98. PDB; 1A9A; 18-NOV-98. Genew; HGNC:2209; COL5A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam; PF01410; COLFI; 1.
Pfam; PF01391; Collagen; 17.
Pfam; PF02210; TSPN; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             4443
1570
11570
11838
1234
2240
2262
2263
3340
3340
3346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1571
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   420
421
570
576
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MIM; 120215; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MIM; 130000; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MOD_RES
MOD_RES
MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MOD_RES
MOD_RES
MOD_RES
MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MOD RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ĝ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ďÖD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ģ'
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ĝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ĝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ĝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ĝ
```

132
Prop. 1885 644 644 PRIDEOUTIATION

```
2240
                                                                                                                                                                                                                 1280
                                                                                                            2183
                                                                                                                                               1243
                                                                                                                                                                                                                                                    2300
                                                                                                                                                                                                                                                                                       1314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                2473
----GPAGEKGAPGEKGPQGPAGRDGLQGPVGLPGPAGPVGPPGEDGD 1157
                                    KSHL-------EGELRPKQPGPVKLGGEAAHLPHLRPLPESQPSSSPLLQTAPGVK 2127
                                                                                                                                                                                                                                                                                                                                                           ------KGSPGPVGFPGDP 1347
                                                                                                                                                                                                                                                                                                                                                                                        2357 WEESPPLSANAFNPLNASASLPAAMPITAADGRSDHTLITSPGGGGKAKVSGRPSSRKAKS 2416
                                                                                                        GHORVVTLAQHISEVITQDYTRHHPQQLSAPLPAPLYSFPGASCPVLDLRRPPSDL----
                                                                                                                                                                                                                                                                                        ---GPKGERGEK----GES
                                                                       1158 KGEIGEPGOKGSKGDKGEOGP--PGPTGPOG-----PIGOPGPSGA---DGEPGPR
                                                                                                                                           1204 GOOGL--FGOKGDE----GPRGFP---GPPGPVGLOGLPG------PPGEKGETG
                                                                                                                                                                            ---YLPPPDHGAPARGSPHSEGGKRSPEPNKTSVLGGGEDGIEPVSPPEGMTEPGHSRSA
                                                                                                                                                                                                               1244 DVGQMGPP-----GPPGPRGPSGAP-----GADG--PQGPPGGIGNPG----A
                                                                                                                                                                                                                                                VYPLLYRDGEQTEPSRMGSKSPGNTSQPPAFFSKLTESNSAMVKSKKQEINKKLNTHNRN
                                                                                                                                                                                                                                                                                                                      2301 EPEYNISOPGTEIFNMPAITGTGLMTYRSQAVQEHASTNMGLEAIIRKALMGKY----DQ
                                                                                                                                                                                                                                                                                                                                                                                                                         1348 GPPGEPGPAGQDGPPGDKGD------DGEPGQT-GSPGPTGEPGPSGPPGKRGPPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                              PA-PGLASGDRPPSVSSVHSEGDCNRRTPLTNRVWEDRPSSAG--STPFPYNPLIMRLQA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1397 PAGPEGRQGEK-GAKGEAGLEGPPGKTGPIGPQGAPGKPGPDGLRGIPGPVG----EQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ĕ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       "The human COL11A2 gene structure indicates that the gene has not evolved with the genes for the major fibrillar collagens."; J. Biol. Chem. 270:22873-22881(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Craniata; Vertebrata; Euteleostomi; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       From
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE OF 730-1690 FROM N.A.
MEDLINE-89340485; PubMed=2760050;
Kimura T., Cheah K.S.E., Chan S.D.H., Lui V.C.H., Mattei M.-G.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Zhidkova N.I., Brewton R.G., Mayne R.; "Molecular cloning of PARP (proline/arginine-rich protein) from cartilage and subsequent demonstration that PARP is a fragment NH2-terminal domain of the collagen alpha 2(XI) chain."; FEBS Lett. 326:25-28(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE=96032717; PubMed=7559422;
Vuoristo M.M., Pihlajamaa T., Vandenberg P., Prockop D.J.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Tubby B.;
Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q13272; Q13273; Q99866; Q9UIP9;
                                                                                                                                                                                                                                                                                                                                                         1315 GPSGAAGPPGPK--GPPGDDGP------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      P13942; Q07751; Q13271; Q13272; Q13273; Q9986 01-JMN-1990 (Rel. 13, Created) 16-OCT-2001 (Rel. 40, Last sequence update) Collagen alpha 2(XI) chain precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GVMASPPPPGLPAGSGPLAGPHHAWDEEPK 2503
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GLPGSPGPPGPMGPPGLPGLKGDSGPK 1479
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1736
                                                                                                                                                                                                                                                                                   1281 V-----GEKGEPGEAGEPGPSGRSGPP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TISSUE-Cartilage;
MEDLINE-93314796; PubMed-8325374;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Eukaryota; Metazoa; Chordata;
Mammalia; Eutheria; Primates;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE OF 59-807 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Homo sapiens (Human)
   1109 POGPP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ala-Kokko L.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HUMAN
                                    2079
                                                                                                                                                                              2184
                                                                                                                                                                                                                                                                                                                                                                                                                                                              2417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1450
                                                                                                        2128
                                                                                                                                                                                                                                                2241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2474
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COLLIA2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Tubby
                                                                  셤
                                  ò
                                                                                                        ઠે
                                                                                                                                    셤
                                                                                                                                                                          à
                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                       .<u>Ş</u>
                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                   엄
                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                         유
                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
```

```
original
                        of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REVIEW ON VARIANTS.
MEDLINE=97255959; PubMed=9101290;
Kuivaniemi H., Tromp G., Prockop D.J.;
"Mutations in fibrillar collagens (types I, II, III, and XI), fibrillassociated collagen (type IX), and network-forming collagen (type X) cause a spectrum of diseases of bone, cartilage, and blood vessels.";
Hum. Mutat. 9:300-315(1997).
der Rest M., Ono K., Solomon E., Ninomiya Y., Olsen B.R.;
human alpha 2(XI) collagen (COL11A2) chain. Molecular cloning of
and genomic DNA reveals characteristics of a fibrillar collagen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sirko-Osadsa D.A., Murray M.A., Scott J.A., Lavery M.A., Warman M.L.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         "Stickler syndrome without eye involvement is caused by mutations in COL11A2, the gene encoding the alpha-2(XI) chain of type XI collagen.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE=95163096; PubMed=7859284; Vikkula M., Marinan B.C.M., Lui V.C.H., Zhidkova N.I., Tiller G.E., Vikkula M., Marinan B.C.M., Lui V.C.H., Zhidkova N.I., Tiller G.E., Van Goldring M.B., van Beersum S.E.C., de Waal Malefijt M.C., van den Hoogen F.H.J., Ropers H.-H., Mayne R., Cheah K.S.E., Olsen B.R., Warman M.L., Brunner H.G.; Mathologian dominant and recessive osteochondrodysplasias associated with the COLI1A2 locus."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Harata S.,
                                                                                                                        SEQUENCE OF 1-537 FROM N.A.

MEDLINE=96435918; PubMed=8838804;
Lui V.C., Ng L.J., Sat E.W., Cheah K.S.;
The human alpha 2 (XI) collagen gene (COL11A2): completion of codi information, identification of the promoter sequence, and precise localization within the major histocompatibility complex reveal overlap with the KES gene.";
Genomics 32:401-412(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Melkoniemi M., Brunner H.G., Manouvrier S., Hennekam R., Superti-Furga A., Kaeaeriaeinen H., Pauli R.M., van Essen T., Warman M.L., Bonaventure J., Miny P., Ala-Kokko L.; Altosomal recessive disorder otospondylomegaepiphyseal dysplasia associated with loss-of-function mutations in the COLIIA2 gene."; Am. J. Hum. Genet. 66:368-377(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE=99021942; PubMed=9805126;
Pihlajamaa T., Prockop D.J., Faber J., Winterpacht A., Zabel B., Aidedon A., Wiesbauer P., Spranger J., Ala-Kokko L.;
Gledion A., Wiesbauer P., Spranger J., Ala-Kokko L.;
"Heterozygous glycine substitution in the Collal gene in the orig patient with the Weissenbacher-Zweymueller syndrome demonstrates identity with heterozygous OSMED (nonocular Stickler syndrome).";
Am. J. Med. Genet. 80:115-120(1998).
                                                                                                                                                                                                                                                                                                                                       MEDLINE=59238468; PubMed=7721876; Zhidkova N.I., Justice S.K., Mayne R.; Alternative mRNA processing occurs in the variable region of pro-alpha 1(XI) and pro-alpha 2(XI) collagen chains."; J. Biol. Chem. 270:9486-9493(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VARIANTS GLY-593; LYS-824; LEU-879; THR-1316 AND GLN-1600.
MEDLINE=98254467; PubMed=9585596;
Koga H., Sakou T., Taketomi B., Hayashi K., Numasawa T., Harata Yone K., Matsunaga S., Otterud B., Inoue I., Leppert M.;
"Genetic mapping of ossification of the posterior longitudinal ligament of the spine.";
"Am. J. Hum. Genet. 62:1460-1467(1998).
                                                                th differences in genomic organization."; Biol. Chem. 264:13910-13916(1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE=20143361; PubMed=10677296;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VARIANT STL3 940-GLY--PRO-948 DEL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE=98165506; PubMed=9506662;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VARIANT OSMED ARG-661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          80:431-437 (1995)
                                                                                                                                                                                                                                                                                                                       ALTERNATIVE SPLICING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VARIANT WZS GLU-955.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Robin N.H.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                    [7]
DISEASE.
                                                                  with
                                                                                  J.
[5]
                                                                                                                                                                                                                                                                                                     9
```

us-09-522-753-5.rsp

\$

```
DISEASE: Defects in COLILA2 are the cause of Stickler syndrome type 3 (STL3) [MIM:184840]. STL3 is an autosomal dominant disorder characterized by oro-facial, auditory and skeletal manifestations,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                beginning
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               such as midfacial hypoplasia, cleft palate, osteoarthritis, and sensorineural hearing loss. Differently from Stickler syndrome type 1 and 2, no coular involvement is observed. This disorder is also referred to as Stickler-like syndrome or non-ocular Stickler
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DISEASE. Defects in COL11A2 are the cause of autosomal dominant monsyndromic sensorineural deafness type 13 (DFNA13) [MM:601868] Affected individuals experience progressive hearing loss beginnin in the second to fourth decades, eventually making use of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DISEASE: Defects in COL11A2 are the cause of autosomal recessive otospondylomegaepiphyseal dysplasia (OSMED) [MIM:215150], a several dysplasia accompanied by severa hearing loss. The phenotype overlaps that of autosomal dominant skeletal disorders (Stickler and Marshall syndromes) but can be distinguished by disproportionately short limbs and lack of ocular involvement. DISEASE: Defects in COL11A2 are the cause of Weissenbacher Zweymueller syndrome (WZS) [MIM:277610], an autosomal dominant disorder allelic with STL3 and OSMED. WZS is also referred to as
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SIMILARITY: BELONGS TO THE FIBRILLAR CLASS OF COLLAGENS.
SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
DATABASE: NAME-Hereditary hearing loss homepage; NOTE-Gene page;
WWW="http://www.uia.ac.be/dnalab/hhh/".
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VSP_001168, VSP_001169;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       extracellular processing and is subsequently retained in the cartilage matrix from which it can be isolated in significant
                                     VARIANTS DFNA13 GLU-808 AND CYS-1034, AND REVISIONS TO 1031-1032
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PTM: A disulfide-bonded peptide called proline/arginine-rich protein or PARP is released from the amino terminus during extracellular processing and is subsequently retained in the
                                                                                                                                                                                                                                                                                                         Event=Allernative splicing; Named isoforms=8;
Comment=Isoforms lack exons 6, 7 or 8 or a combination of
exons. Experimental confirmation may be lacking for some
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IsoId=P13942-8; Sequence=VSP 001167, VSP 001168, VSP 0011
PTM: Prolines at the third position of the tripeptide repea
unit (G-X-Y) are hydroxylated in some or all of the chains.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             in the second to fourth decades, eventually making use of amplification mandatory.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VSP_001168;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VSP_001169;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VSP_001169;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        IsoId=P13942-6; Sequence=VSP_001167,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    IsoId=P13942-5; Sequence=VSP_001167,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IsoId=P13942-7; Sequence=VSP_001168,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IsoId=P13942-3; Sequence=VSP 001168;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  IsoId=P13942-4; Sequence=VSP_001169;
                                                                                                                                                                                                                                                                                                                                                                                                       IsoId=P13942-1; Sequence=Displayed;
 Pediatr. 132:368-371(1998)
                                                                                                                                                                                                                                                                       of alpha 3(XI)=1(II).
ALTERNATIVE PRODUCTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           heterozygous OSMED.
                                                                                                                                                                                                                                                                                                                                                                                                                                          IsoId=P13942-2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           syndrome
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Name=3
                                                                                                                                                                                                                                                                                                                                                                                      Name=1
                                                                                                                                                                                                                                                                                                                                                                                                                         Name=2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ---
```

```
97;
               There are no restrictions on its 
ng as its content is in no way 
red. Usage by and for commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PPPQNLQPESD--APQQPGSSPRGKSRSPAPPADK-EAFAAEAQKLPGDPPC-----W 1049
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEASVLRGTALGSVPGGSITKG 1211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          816
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           468
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PR-PPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEHAKAPVG----PVTMG 1151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         658
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1212 IPSTRVPSDSAITYRGSITHGTPADVL-YKGTITRIIGEDSPSRL----DRGREDSLP- 1264
                                                           entities requires a license agreement (See http://www.isb-sib.ch/announce,
                                                                                                                                                                                                                                                                                                                                                                                                                  701
                                                                                                                                                                                                                                                                                                                                                                                                                                                191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               756
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     863
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                897
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             939
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            613
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      691
 EMBL outstation
                                                                                                                                                                                                                                                                                                                                                                                    OGVROLGLELGRPVRFLYEDQTGRPQPPSQPVFRGLSLADGKWHRVAVAVKGQSVTLIVD
                                                                                                                                                                                                                                                                                                                                                                                                                                                -----KRVTRPLPRSAR-----PVLDTHGVIIF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EAAKDTGQNGPKPPATLGADGPPPGPPTPPRRTSRAPIEPTPASEATGAPTPPPAPSPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        411 GPPGIQGNPGPVGDPGERGPPGRAGLPGSDGAPGPPGTSLMLPFR--FGSGGGDKGPVVA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRPSLLTPTGDPRANASPQKPLDLKQLKQRAAAIPPIQVTKVHEPPREDAAPTKPAPPAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AQ------PPGPMGYTGRPGPLGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PGSPGLKGESGDLGPOGP-RGPOGLTGSLGKAGRRGRAGADGARGMPGDPGVKGDRGFDG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               --FPVPPREVIKASPHAPDPSAFSYAPPGHPLPLGLHDTARPVL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              563 LPGLPGEKGHRGDTGAQGLPGPPGE---DGERGDDGEIGPRGLPGESGPRGL-----LG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PKGPPGIPGPPG-VRGMDGPQGPKGSLG-------PQGB--PGPPGQGTPGTQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LPGP------GANGPHGEKGPQGKPGLP------GMPGSDGPPG
                                                                                                                                                                                                                                                                                                                                                       OSAELASMELNESSRWTEEEM-----ETAKKGLLEHGRNWSAIARMVGSKTVS---O
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ----DEEMEASGVSGNEEEMVEEAEALHASGNEVPRGECSGPATVNNSSDTESIPSPHT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ----RAQRSPQQQPSRLHRPQNQEPQSQPTESLYYDYEPPYYDVMTTGTTPDYQDPTPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPPVVPKEEKEEETAAAPPVEEGEE--OKPPAAEELAVD-----TGKAEEPVK----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ---------GGSGRATTAKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLS
                                                                                                                                                                                                                                                                                                                                                                                                                  CKNFYFNYKKRQNLDEILQQHKLKMEKERNARRKKKKAPAAASEEAAFPPVVE----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GARILDEEVPEGDV--QELAIVPGVQAAYES-CEQKELECEG----GQRERPQNQQPH-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GDDYREETELGPALSAETAHSGAAAHGPRGLKGEKGEPAVLEPGMLVEGPPGPEGPAGLI
                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                         629
                                                                                                                                                                                                                                                                                            Length 1736;
                                                                                                                                                                                                                                                                                                                       Indels
Swiss Institute of Bioinformatics and the
                                                                                                                                                                                                                                                                                         Score 350.5; DB 1;
Pred. No. 0.00015;
8; Mismatches 704;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -SECTEEAEEGPAKGKDAEAAEATAEG--ALKAEKKE-
              the Buropean Bioinformatics Institute. Then use by non-profit institutions as long amodified and this statement is not removed.
                                                                         email to license@isb-sib.ch)
                                                                                                                                                 ALO31228; CAA20240.1; -...
118987; AAA35498.1; -...
1041069; AAAC17464.1; -...
1041065; AAC17464.1; -...
1041065; AAC17464.1; JOINED.
1041066; AAC17464.1; JOINED.
                                                                                                                                                                                                                                                                                                                       Conservative 138;
                                                                                                                                                                                                                                                                                          21.0%;
                                                                                                      EMBL; U32169; AAC50214.1; -. EMBL; U32169; AAC50213.1; -. EMBL; U32169; AAC50215.1; -.
                                                                                                                                                                                                                                                            COLLIA2
                                                                                                                                                                                                                                                                                                          Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                  CK----
                                                                                                                                                                                                                                                            HGNC: 2187;
 the
                                                                                                                                                                                                                                                                                       Query Match
Best Local Simi]
Matches 400; (
                                                                           or send an
                                                                                                                                                                                                                                                                                                                                                       900
                                                                                                                                                                                                                                                                                                                                                                                     107
                                                                                                                                                                                                                                                                                                                                                                                                                  649
                                                                                                                                                                                                                                                                                                                                                                                                                                                167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               702
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          757
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      817
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  864
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              868
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         940
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1050
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1097
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       629
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        243
   Setween
                                                                                                                                                                                                                                                             Genew;
                                                                                                                                                                                 EMBL;
EMBL;
                                                                                                                                                                                                                            EMBL;
EMBL;
                                                                                                                                                    EMBL;
                                                                                                                                                                    EMBL;
                                                                                                                                                                                                                 SMBL;
셤
                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         용
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   용
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      . අ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                        ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              용
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ਨੇ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              . 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
```

= = ::

This SWISS-PROT entry is copyright. It is produced through a collaboration

<i>&</i> 8	55	DT DE ACCOUNTS OF THE ACCOUNTS		R R R R R R R R R R R R R R R R R R R												RT re
692 HPGKEGPPGTKGNQGPSGPQGPLGYPGP-RGVKGVDGIRGLKGHKGEKG-EDGFPG 745 1265 -KGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETAAP	MMEGRVGRAISSASIEGLMGRAIPPERHSPHHLKEQHHIRGSI	BS9 GAKGTSGGDGPHGPPGERGLPGPQGPNGFPGPKGPFGPPGPPG 903 1396 AYKTQALGPLKLKPAHEGLVATVKEAGRSIHEIPREELRHTPELPLAPRPLK 1447 1396 AYKTQ	1448 EGSITQGTPLKYDTGASTTGSKKHDVRSLIGSPGRTFPPVHPLDVWADAR 1497 	1498 ALERACYEESLKSRPGTASSSGGSIARGAPVIVPELGKPROSPLTYEDHGAPFAGHLP 1555 	1556 RGSPVTMREPTPRLOEGSLSSSKASQDRKLTSTPREIAKSPHSTVPEHHPHPISPYEHLL 1615	1616 RGVSGVDLYRSHIPLAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIRGYPD 1675	1676 TAALENRQTIINDYITSQQMHNNTATAMAQRADMLRGLSPRESSLALNYAAGPRGIIDLS 1735	1736 QVPHLPVLVPPTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTHLTKP 1788 1169 ETGDVGPMGPPGPPGPRGPAGPNGADGPQGPPGGVGNLGPPGEKGEP 1215	1789 TITSSSERERDRDRERDRDREREKSILISTITVEHAPIWRPGTEQSSGSSGSGGGG 1846 1216 GESGSPGIQGEPGVKGPRGERGEKGESGQPGEPGPPGAKGPDGDDGPKGNPGPVG 1270	1847 SSSRPASHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGIITAVEPSKPTVLRSTSTSSP 1906	1907 VRPAATFPPATHCPLGGTLDGVYPTLMEPVLLPKEAPRVARPERPRADTGHAFLAKPPAR 1966 1306 GPTGENGPPGPLG	1967 SGLEPASSPSKGSEPRPLVPPVSGHATIARTPAKNLAPHHASPDPPAPPASBSDPHREKT 2026 1320 RGPAGSPGSEGRQGGKGAKGDPGAIGAPGKTGPVGPAGFAGKPGPDGL 1367	2027 QSKPFSI-QELELRSLGYHGSSYSPEGVEPVSPSSPSLTHDKGLPKHLEELDKSHLE 2083 : : :	2084 GELRPKQPGPVKLGGEAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQRVVTLAQHISEVI 2143	2144 TQDYTRHHPQQLSAPLPAPLYSFPGASCPVLDLRRPPSDLYLPPPDHGAPARGSPHSEGG 2203 	2204 KRSPEPNKTSVLGGGEDGIEPVSPPEGMTEPGHSRSAVYPLLYRDGEQTEPSRMGSK 2260
9 & 9	8 8 8	8 & 8	දි පි	\$ g	දු පු	දු දු	දු පු	දු පු	\$ g	\$ A	8 %	o o	දු පු	S 원	상 원	b ox

```
Oy 2261 -------SERNITGORPAPESSKITTENSISAMYKKKOETHK 2222

D 1524 EDEALPPGAARGSPGALERIFGSL.-----DSLREEIED 1556

REBUIT 54

APC-INDAN

C 2561, Olisic2, Olisic3;
D 10-AMX-1992 [Rel. 22, Lest esquence update)
D 15-AMX-1992 [Rel. 22, Lest esquence update)
D 15-AMX-1992 [Rel. 22, Lest esquence update)
D 16-AMX-1992 [Rel. 22, Lest esquence update)
D 20-AMX-1992 [Rel. 22, Lest esquence update)
D 30-AMX-1992 [Rel. 22, Lest esquence update)
D 30-AMX-1992 [Rel. 22, Lest esquence update)
D 40-AMX-1992 [Rel. 22, Lest esquence update)
D 50-AMX-1992 [Rel. 22, Rel. Milbert L. AMX-1992 P. C. Mandallica P. Milbert L. AMX-1992 P. C. Mandallica P. Milbert L. AMX-1992 [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Rel. 22] [Re
```

```
∴ ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE=93250848; PubMed=1338764;
Nagase H., Miyoshi Y., Horii A., Aoki T., Petersen G.M.,
Nagase H., Mahlor E., Ogawa M., Maruyama M., Utsunomiya J.,
Baba S., Nakamura Y.;
"Screening for germ-line mutations in familial adenomatous polyposis
patients: 61 new parients and a summary of 150 unrelated patients.";
Hum. Mutat. 1:467-473(1992).
                                                                                                                                                                                                                                                                                                                                                                                                            Nishisho I., Nakamura Y., Miyoshi Y., Miki Y., Ando H., Horii A., Koyama K., Utsunomiya J., Baba S., Hedge P., Markham A., Krush A.J., Petersen G.M., Hamilton S.R., Nilbert M.C., Levy D.B., Bryan T.M., Preisinger A.C., Smith K.J., Su L.-K., Kinzler K.W., Vogelstein B.; "Mutations of chromosome 5q21 genes in FAP and colorectal cancer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Nakatsuru S., Yanagisawa A., Ichii S., Tahara E., Kato Y.,
Nakamura Y., Horii A.,
"Somatic mutation of the APC gene in gastric cancer: frequent
mutations in very well differentiated adenocarcinoma and signet-ring
                                                                                                                                                                                                       K-RAY CRYSTALLOGRAPHY (1.90 ANGSTROMS) OF 2034-2049 IN COMPLEX WITH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    VARIANT FAP TRP-1348, AND VARIANTS ASP-1118; MET-1292; VAL-1304 AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ٦
:
                                      the
                                                                                       OF 1021-1035 IN COMPLEX WITH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE=91265010; PubMed=1338904;
MEDLINE=91265010; PubMed=1338904;
Miyoshi Y., Nagase H., Ando H., Ichii S., Nakatsuru S., Aoki T.,
Miki Y., Mori T., Nakamura Y.;
"Somatic mutations of the APC gene in colorectal tumore: mutation cluster region in the APC gene.";
                                                                                                              MEDLINE=21564054; PubMed=11707392;
Bklof Spink K., Fridman S.G., Weis W.I.;
"Molecular mechanisms of beta-catenin recognition by adenomatous
polyposis coli revealed by the structure of an APC-beta-catenin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TISSUE-Peripheral blood lymphocytes,
MEDLINE-95134944; PubMed-7833149;
Dobbie 2., Spycher M., Huerliman R., Ammann R., Ammann T., Roth
Mueller A., Mueller H., Scott R.J.;
"Mutational analysis of the first 14 exons of the adenomatous
X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 2-55.
MEDLINE=20384842; PubMed=10926498;
Day C.L., Alber T.;
"Crystal structure of the amino-terminal coiled-coil domain of
                                                                                                                                                                                                                                                                                                                                   Nagase H., Nakamura Y.; "Mutations of the APC (adenomatous polyposis coli) gene."; Hum. Mutat. 2:425-434(1993).
                                                                                                                                                                                                                                           Spink K.E., Polakis P., Weis W.I.; "Structural basis of the axin-adenomatous polyposis
                                                                                       K-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS)
                                                                                                                                                                                                                                MEDLINE=20271867; PubMed=10811618;
                                                                                                                                                                                                                                                                                                                                                                                      VARIANTS FAP.
MEDLINE=91335211; PubMed=1651563;
                                                                                                                                                                                                                                                                                                                        MEDLINE=94154728; PubMed=8111410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE=93244793; PubMed=1338691
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Mol. Genet. 1:229-233 (1992)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Mol. Genet. 1:559-563(1992)
                                                              . Mol. Biol. 301:147-156(2000)
                                                                                                                                                                     complex.";
EMBO J. 20:6203-6212(2001)
                                                                                                                                                                                                                                                                                   EMBO J. 19:2270-2279(2000)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Science 253:665-669(1991).
                                                   tumor suppressor
                                                                                                                                                                                                                                                                                                           REVIEW ON VARIANTS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VARIANT FAP TRP-99.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 carcinoma.
                                                                                                                                                                                                                                                                       interaction.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VARIANTS FAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             patients.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VARIANTS
```

```
92 LPELGKSEMEFIESKRPRLELLPDPLLRPSPLLATGQPAGSEDLT---KDRSL----- 141
                                                                                                                                                                                                    Stella A., Montera M., Resta N., Marchese C., Susca F., Gentile M., Romio L., Pilia S., Prete F., Mareni C., Guanti G.; "Four novel mutations of the APC (adenomatous polyposis coli) gene in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SYSO ASHLSPGSIIOP--ORRRPSILSEPO-------PGNERSQELHLRPESHSY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VARIANTS COLORECTAL CARCINOMA THR-880; ILB-890 AND VAL-1508.
MEDLINE-98080146; PubMed-9419979;
Miyaki M., Nishio J., Konishi M., Kikuchi-Yanoshita R., Tanaka K.,
Muraoka M., Nagato M., Chong J.-M., Koike M., Terada T., Kawahara Y.,
Pukutome A., Tomiyama J., Chuganji Y., Momoi M., Utsunomiya J.;
"Drastic genetic instability of tumors and normal tissues in Turcot
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              740 ANIMSPGSSLPSLHVRKOKALEAELDAOHLSETFDNIDNLSPKASHRSKORHKOSLYGDY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     study of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ά
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              F., Gentile M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TISSUE=Peripheral blood;
MEDLINE=98393112; PubMed=9724771;
Frayling I.M., Beck N.E., Ilyas M., Dove-Edwin I., Goodman P.,
Frayling I.M., Beck N.E., Ilyas M., Dove-Edwin I., Goodman P.,
Pack K., Bell J.A., Williams C.B., Hodgson S.V., Thomas H.J.W.,
Talbot I.C., Bodmer W.F., Tomlinson I.P.M.;
"The APC variants I1307K and B1317Q are associated with colorectal
tumors, but not always with a family history,",
Proc. Natl. Acad. Sci. U.S.A. 95:10722-10727(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WEDLINE=98400248; PubMed=9731522;
MEDLINE=98400248; PubMed=9731522;
MEDLINE=98400248; PubMed=9731522;
Medston M., Nathanson K.L., Yuan Z.Q., Neuhausen S.L., Satagopan , Wong N., Yang D., Nafa D., Abrahamson J., Ozcelik H.,
Antin-Ozerkis D., Andrulis I., Daly M., Pinsky L., Schrag D.,
Antin-Ozerkis S., Kaback M., King M.-C., Woodage T., Brody L.C.,
Godlinger S., Kaback M., Fing M.-C., Woodage T., Brody L.C.,
"The APC I1307K allele and breast cancer risk.";
Nat. Genet. 20:13-14(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           familial adenomatous polyposis: 67 germline mutations identified DGGE, PTT, and southern analysis."; Hum. Mutat. 9:7-16(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VARIANT LYS-1307.
MEDLINE-98400259; PubMed-9731533;
Woodage T., King S.M., Wacholder S., Hartge P., Struewing J.P.,
Moodams M., Laken S.J., Tucker M.A., Brody L.C.;
"The APC I1307k allele and cancer risk in a community-based stuc
Ashkenazi Jews.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels 821;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             van der Luijt R.B., Meera Khan P., Vasen H.F.A., Tops C.M.J., van Leeuwen-Cornelisse I.S.J., Wijnen J.T., van der Klift H.M., Plug R.J., Grifficen G., Fodde R.; Molecular analysis of the APC gene in 105 Dutch kindreds with
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 2.7%; Score 350.5; DB 1; Length 2843; Best Local Similarity 17.8%; Pred. No. 0.00023; Matches 473; Conservative 357; Mismatches 1007; Indels 821;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Stella A., Montera M., Resta N., Marchese C., Susca
Romio L., Pilia S., Prete F., Mareni C., Guanti G.;
Hum. Mol. Genet. 3:1918-1918(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           :-
polyposis coli (APC) gene.";
Eur. J. Cancer 30A:1709-1713(1994).
                                                                                                                                                                                                                                                                                                                                                                      Mol. Genet. 3:1687-1688(1994).
                                                                                                                       VARIANT FAP GLY-722.
MEDLINE=95135430; PubMed=7833931;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VARIANT FAP ILE-171.
MEDLINE=97144176; PubMed=8990002;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VARIANTS LYS-1307 AND GLN-1317.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Oncogene 15:2877-2881(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Nat. Genet. 20:62-65(1998).
                                                                                                                                                                                                                                                                                                                                    FAP patients."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          syndrome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                            ERRATUM
```

	1304 TAAPKRTYDMMEGRV 1846YTPIEGTP 1356 TQGIPRSYVEAQEDY 1895 TSHTELTSNQQSANK 1401 ALGPLKLKPAHEGLV 1955 AIENTPVCFSHNSL 1955 AIENTPVCFSHNSL 1453 QGTPLKYDTGASTTG 1453 QGTPLKYDTGASTTG 1615 EDTPVCFSRNSSLSS 1483 TFPPVHPLDVMADAR	2075	2214KQ 1764 APQPFSSRHSSSPLS 1823RNSSSSTSPVS; 1821VEHADIWRP 1274 AKPSVKSELSPVARQ 1875 PSVLHNTGMKGITTA 2330SPGRNGI: 1931TLMEPVLLPREA 1931TLMEPVLLPREA 1931PRPLVPPVSGH 2374 SQNITKQTGLSKNA 1981PRPLVPPVSGH 2429 SDRSERPVLVRQSTF 2036 LELRSLGYHGSSYSP)	2488 LPDMSLSTH-SSVQA
8 8 8 8 8	8 8 8 8 8 8 8 8	8 8 8 8 8 8 8	8 8 8 8 8 8 8 8 8 8 8	Q
800 VFDTNRHDDNRSDNFNTGNMTVLSPYLNTTVLPSSSSSRGSLDSSRSEKDRSLERERGIG 859 142 TGKLEPVSPPSPPHTDFELELVPPRLSKEELIQNMDRVDREITMVEQQISKLKKKQQQLE 201 860 LGNYHPATEN	PRRRAKESKVREYYEKOFP-EIRKORELOERMOSRVGGRGGGLSMGARSE	DYSLKYATDI PSSGKQSFSFSKSSGQSSKTEHMSSSSENTSTPSSNAKRQNQLHPSSA NEANSEEAITPQGSAELASMELNESSRWTEEEMETAKKGLLEHGRNWSAIARMVGSKTVS S.		1089 HUIAKKVLKKEFILISNEFELISSANAFSVLEKQIGAISQGMSVQLAVKISEHAKAFVGEV 1148
8 6 6 6 6 6	6 8 8 8 8 8 8 8	6 8 6 8 6 8 6 8 6	6 8 6 8 6 8 6 8 6 8 6 8 6	3 8

1149	TMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEASVLRGTALGSV 1203 ::::::::::::::::::::::::::::::::::::
1204	PGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYKGTITRI 124
1733	
1247	IGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHE 1303
1304	TAAPKRTYDMMEGRVGRAISSASIEGLMGRAIPPERHSPHHLKEQHHIRGSI 1355
1356	TOGIPR TSHTEL
1401	ALGPLKLKPAHEGLVATVKEAGRSIHEIPREELRHTPELPLAPRPLKEGSIT 1452
1453	QGTPLKXDTGASTTGSKHDVRSLIGSPGR 1482 : : :
1483	TFPPVHPLDVMADARALERACYEESLKSRPGTASSSGGSIARGAPVIVPELGKPRQSPLT 1542
1543	YEDHGAPFAGHL.PR
1585	LTSTPREIAKSPHSTVPEHHPHPISPYEHLLRGV-SGVDLYRSHIPLAFDPTSIPRGIPL 1643
1644	DAO :
1704	AQRADMLRGLSPRESSLALNYAAGPRGIIDLSQVPHLPVLVPPTPGTPATAMDRLAYLPT 1763 :
1764	APQPFSSRHSS : : RNSSSST
1821	VEHAPIWRPGTEQSSGSSGGGGGSSSRPASHSHAHQHSPISPRTQDALQQR 1874 :
1875	PSVLHNTGWKGIITAVEPSKPTVLRSTSSPVRPAATFPPATHCPLGGTLDGVYP 1930
1931	TIMEFVILPKEAPRVARPERPRADTGHAFLAKPPARSGLEPASSBSKGSE 1980
1981	PRPLVPPVSGHATIARTP :: SDRSERPVLVRQSTFIKEAPSP
2036	LELRSLGYHGSSYSPEGVEPVSPYSSPSLTHDKGLPRHLEELDKSHLEGELRPKQPGPVK 2095

```
2806
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                       2764
LGG--EAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQRVVTLAQHISEVITQDYTRHHPQ 2153
                                        RSGTWKREHSKH----SSSLPRVSTWRRTGSSSSILSASSESSEKAKSEDEKH---
                                                                                  2154 QLSAPLPAPLYSFPGASCPVLDLRRPPSDLYLPPPDHGAPARGSPHSEGGKRSPEPNKTS
                                                                                                                          ------VNSISGTK-----QSKENQVSAKGTWRKIKENEFSPTNSTS
                                                                                                                                                                       --VLGGGEDGIE-----PVSPPEGMTEPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNT
                                                                                                                                                                                                             2628 QTVSSGATNGAESKTLIYQMAPAVSKTEDVWVRIEDCPI------NNPRSGRSPTGNT
                                                                                                                                                                                                                                                       SQPPAFFSKLTESNSAMVKSKKQEINKKLNTHNRNEPEYNISQPGTEIFNMPAITGTGLM
                                                                                                                                                                                                                                                                                                  ---GNGSV
                                                                                                                                                                                                                                                                                                                                            2326 TYRSOAVOEHASTNMGLEAIIRKALMGKYDQWEESPPLSANAFNPLNASASLPAAMPITA
                                                                                                                                                                                                                                                                                                                                                                                    PMRTVGLENRLNSFIQVDAPDQKGTEIKPGONNPVPVSETN-----ESSIVERTPFSS
                                                                                                                                                                                                                                                                                                                                                                                                                                 ADGRSDHTLTSPGGGGKAKVSG---RPSSRKAKSPAPGLASGDRPPSVSSVHSEGDCNRR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SSS-SKH--SSPSGTVAARVTPFNYNPSPRKSSAD----STSARPSQIP-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sherrington R., Rogaever. Liang Y., Rogaeva B.A., Levesque G., Ikeda M., Chi H., Lin C., Li G., Holman K., Tsuda T., Mar L., Foncin J.-F., Bruni A.C., Montesi M.P., Sorbi S., Rainero I., Pinessi L., Nee L., Chumakov I., Pollen D., Brookes A., Sanseau P., Polinsky R.J., Wasco W., da Sliva H.A.R., Haines J.L., Rommens J.M., St George-Hyslop P.H.;

Rommens J.M., St George-Hyslop P.H.;

"Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease.";

Nature 375:754-760(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Rowen L., Madan A., Qin S., Abbasi N., Baradarani L., Birditt Bloom S., Dors M., Dickhoff R., Fleetwood P., Harrison G., Jam Kaur A., Madan A., Owen M.P., Ratcliffe A., Shaffer T., Hood L. "Sequencing of human chromosome 14q24.3 region."; Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     re 375:754-760(1995).
SUBCELLULAR LOCATION: Nuclear (Potential).
CAUTION: Ref.2 sequence differs from that shown due to
frameshift in position 1661.
                                                                                                                                                                                                                                                                                                  --PP-----VIDSVSEKANPNIKDSKDNQAKQNV--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    P49750; P49752; Q9P1V7;
01-0CT-1996 (Rel. 34, Created)
16-0CT-2001 (Rel. 40, Last sequence update)
16-0CT-2001 (Rel. 40, Last annotation update)
Nuclear protein ZAP3 (ZAP113).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1822 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE OF 539-847 AND 1397-1822 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE=95319502; PubMed=7596406;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2443 TPLTNRVWEDRPSSAGST 2460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2807 TPVNNNT-KKRDSKTDST 2823
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FISSUE=Brain;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ZAP3 HUMAN
2096
                                        2543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2765
                                                                                                                          2592
                                                                                                                                                                     2214
                                                                                                                                                                                                                                                     2266
                                                                                                                                                                                                                                                                                                  2680
                                                                                                                                                                                                                                                                                                                                                                                       2712
                                                                                                                                                                                                                                                                                                                                                                                                                                 2386
                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                  Š
                                                                                                               g
                                                                                                                                                                   Š
                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
```

```
91;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QRAAAIPPIQVTKVH-------EPPREDA-----APTKPAPPAPP- 1001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1081 GHPLPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLH----- 1134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1169 KQ-EQL-----SPRGQAGPPESLGVPTAQEASVLRGT--ALGSVPGGSITKG--I 1212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PPERHSPHHLKEQHHIRGSITQGIPRSYVEAQEDYLRREAKLLKREGTPPP---PPPSR- 1391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             802
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               737
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             203 ----PTPS------YSSSSSSSSSSSSSSYLSHSGSYLPSSQASPSRPSQGHSKSQ 243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 910 --GAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRANASPQKPLDLKQLK 967
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 244 ILAPPPSAPPGNKTTVQQEPLESGAKNK--STEQQQAAPEPDP----STMTPQEQQYW 297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TOIDKAQAVTOPVPLANKPVPAQSTFPSKT-GGMEG--GTAVATSSLTADNDFKPVGIGL 794
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     87 PPPPVMPGGGYGDWQPPPPPPPPPPGPALSYQKQQQYKHQMLHHQRDGPPGLVPMELESP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              298 YRQHLLSLQQRTKVHLPGHKKGPVVAKDTPEPVKEEVTVPATSQVPESPSSEEPPLPPPN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1002 ---PQNLQPESDAPQQPGSSPRGKSRSPA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----PPADKEAFAAEAQKLPGDPPCWTSGLPFPVPPREVIKASPHAPDPSAFSYAPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EKPRP-----SAGPSEGSAPP---TIYHPPLQ-----SAGPSEQVNSKAPLS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          :|| | : | | | : | | | : | | : | | : | | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              629 PSTSLSPRQSGPQWKGPKPAFGQQHQQQPKSQAEPLSGNKEPLADTSSNQQKNFKMQSAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HGTPADVLYKGTITRIIGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSV----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ---RRTSRAPIEPTPASEA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TGAPTPPPA----PPSPSAPPPVVPKEEKEETAA---APPVEEGEEQKPPAAEELAVDTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               135 ---VPYS-----EHAKAPVGPVTMGLPLPMDPKKLAP------FSGV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TQCSKEDGRSSSGP-----PHETAAPKRTYDMMEGRVGRAISSASIEG-----LMGRAI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       857 KAEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSS----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           358 ЕВУРРРГРРЕВРОЅБОРЕБОАКТКОГОАЛАННООНООНКУСРОУОСІМОКНТОГООІГО
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            147 PESPPVPPGSYMPPSQSYMPPPQPPSYYPPTSSQPYLPPAQPSPSQSPPSQSYLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 1822;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                             CRC64;
                                                                                                                                                                                                                                                                                               P -> S (IN REF. 2).
T -> I (IN REF. 2).
K -> E (IN REF. 2).
MW, 8E6CB83FE540C7D2
                                                                                                                                                                                                                                                                                                                                                                                                                                                   2.6%; Score 349.5; DB 1;
ilarity 20.2%; Pred. No. 0.00016;
Conservative 157; Mismatches 642;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PKPPATLG---ADGPPGPPTPP--------
send an email to license@isb-sib.ch)
                                                                                                                                                                               PRO-RICH.
GLN-RICH.
ARG-RICH.
ARG-RICH.
                                                                                       EMBL; L40400; AAC42008.1; ALT_FRAME.
EMBL; L40400; AAC42006.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PSTRV-PSDSAITYRG-----
                                                                                                                                                                                                                                                                                                                                                                                          AA; 204947
                                                        EMBL; AC007956; AAF61275.1;
                                                                                                                                                                                                              430
1209
1577
621
1404
1821
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Similarity
                                                                                                                                                                                                                                       807
1488
                                                                                                                                                         protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              datches 370;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            469 1
                                                                                                                                               Nuclear p
DOMAIN
DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    803
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1028
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  968
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1287
                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query
   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  윰: &
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
```

ΩI	A C C C	OSUB	R O O C	RX RA RA	RI RI RN RD	R R R R R R R R R R R R R R R R R R R	RA RA RA	RA RA RA	RA RA RA RA	RA RA RT	RT RN RP	RX RA RT	RL RP RX	RA RT RT	RP RP RX	RA RA RA	RA RA RA
795 PHSENNQDKGLPRPDNRDNRLEGNRGNSSSYRGPGQSRM 833	1392DLTEAYKTQALGPLKLKPAHEGLVATVKEAGRSIHEIPREELRHTPEL	1440 PLAPRPLKEGSITQGTPLKYDTGASTTGSKKHDV-RSLIGSPGRTFPPVHPLDVMADARA 1498	LERACYEESLKSRPGTASSGGSIARGAPVIVPELGKPRQSPLTYEDHGAPFAGHL	PRGSPVTMREPTPRLQEGSLSSSKASQDR-KLTSTPREIAKSPHSTVPEHHPH-PISPYE	HILRGVSGVDLYRSHIPLAFDPTSIP	HLAPNPTYPHLYPPYLIRGYPDTAALENRQTIINDYTTSQQMHHNTATAMAQRADML	RGLSPRESSLALNYAAGPRGIIDLSQVPHLPVLVPPTPGTPATAMDR	1758 LAYLPTAPQPFSSRHS-SSPLSPGGPTHLTKPTTTSSSBRERDRDRERDRDR 1808 	1809 EREKSILITSTTTVEHAPIWRPGTEQSSGSSGGGGGSSSRPASHSHAHQHSPI 1863 :	1864 SPRIQDALQORPSVLHNTGMKGIITAVEPSKPTVLRSTSTSSPVRPAATFPPATH 1918 	1919 CPLGGTLDGVYPTLMEPVLLPKEAPRVARPERPRADTGHAFLAKPPARSGLEPASSPSKG 1978 	1979 SEPRPLVPPVSGHATIART 1997 	1998 PAKNLAPHHASPDPPAPASASDPHREKTQSKPFSIQELELRSLGY-HGSSYSPEGVEPV 2056	2057SPVSSPSLTH 2066	- DKGLPKHLEELDKSHLEGELRPKQPGPVKLGGERAHLPHLRPLPESQPSSSPLLQT 	2123 APGVKGHQRVVTLAQHISEVITQDYTRHHPQQLSAPLPAPLYSFPGASCPVLDLRRPPS- 2181	2182DLYLPPPDHGAPARGSPH 2199
qa	ç, Bb	, da		٠ و د			Oy 1	9 9 G	. Oy da	99 1	99 Dp	9y 1	95 1	Öy.		Oy 2 Db 1	Oy 2 Db 1

RESULT 56 CA13_MOUSE

DOLINOUSE STRUMENT: PRT: 1464 AA. POBLING: STRUMENT: PRT: 1464 AA. POBLING: STRUMENT: PRESENT STRUMENT

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
                                                                                                                                                                                                                                                                                                                                                 PTM: O-linked glycan consists of a Glc-Gal disaccharide bound to the oxygen atom of a post-translationally added hydroxyl group (By
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NONHELICAL REGION (N-TERMINAL).
TRIPLE-HELICAL REGION.
NONHELICAL REGION (C-TERMINAL).
O-LINKED (GAL. . .) (BY SIMILARITY).
HYDROXYLATION (BY SIMILARITY).
                                                                      Hayashizaki Y.; "Punction of a full-length mouse cDNA collection."; Nature 409:685-690(2001).
                                                                                                                                                                          Metsaeranta M., Toman D., de Crombrugghe B., Vuorio E.;
"Specific hybridization probes for mouse type I, II, III and IX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PROSITE; PS01208; VWPC_1; 1.
PROSITE; PS50184; VWPC_2; 1.
Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AMINO-TERMINAL PROPEPTIDE.
COLLAGEN ALPHA 1 (III) CHAIN.
CARBOXYL-TERMINAL PROPEPTIDE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BY SIMILARITY.
                                                                                                                                                                                                                                                                                                                                                                                similarity).
SIMILARITY: Contains 1 VWFC domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PIR, SS9856; SS9856.
MGD; MGI:88453; Col3al.
InterPro; IPR008161; Collagen.
InterPro; IPR008161; Collagen.
InterPro; IPR001885; Fib_collagen.
InterPro; IPR001885; Fib_collagen.
InterPro; IPR001897; VWF_C.
Pfam; PF01410; CoLFI; 1.
ProDom; PD000007; CIG helix; 1.
ProDom; PD000007; CIG helix; 1.
ProDom; P0002078; Fib_collagen_C; 1.
SMART; SM00218; Fib_collagen_C; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; M18933; AAA37338.1; -.
EMBL; K03037; -; NOT ANNOTATED_CDS.
EMBL; AK019448; BAB31724.1; -.
EMBL; K57983; CAA41048.1; -.
PIR; A27353; A27353.
                                                                                                                                            STRAIN=C57BL/6;
MEDLINE=91274355; PubMed=2054384;
                                                                                                                                SEQUENCE OF 1442-1464 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; X52046; CAA36279.1; -.
EMBL; BC043089; AAH43089.1; -.
EMBL; BC058724; AAH58724.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Collagen; Signal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            23
154
1203
1464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1464
262
262
283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Glycoprotein,
SIGNAL
                                                                                                                                                                                                                                                                                                                                      chains.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
MOD_RES
MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
PROPEP
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
```

```
1138 SEHAKAPVGPV-ŢMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEASVLR 1196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               197 GTALGSVPGGSITKGIPSTRVPSDSAITYRGSI-THGTPADVLYKGTITRIIGEDSPSRL 1255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1482 RTFPPVHPLDVMADARALERACYEESLKSRPGTASSSGGSIARGAPVIVPELGK-PRQSP 1540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1019 -PRGKSRSPAPPADKEAFAAEAQKLPGDPPCWTSGLPFPVPPREVIKASPHAPDPSAFSY 1077
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1541 LTYEDHGAPFAGHLP--RGSPVTMREPTPRLQEGSLSSSKASQDRKLTSTPREIAKSPHS 1598
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1078 APPGHPLPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHVPY 1137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1316 GRVGRAISSASIEGLMGRAIPPERHSPHHLKEQHHIRG----SITQGIPRSYVEAQEDYL 1371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1432 BLRHTPELPLAPRPLKEGSITQGTP-----LKYDTGASTTGSKKHDVRSLIGSPG 1481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1372 RREAKLLKREGTPPPPPPPSRDLTEAYKTQALGPLKLKPAHEGLVATVKEAGRSIHEIPRE 1431
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        712
                                                                                                                                                                                                                        767
                                                                                                                                                                                                                                                           108
                                                                                                                                                                                                                                                                                                                      845
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        906 AKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPR-ANASPQKPLDLK 964
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   410
                                                                                                                                                                                                                                               DIICDEEPLDCPNPEIPFGECCAICPQPSTPAPV-----LPDGH-------GPQGP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PGA-----LGPA------GPA-GKDGE------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1256 DRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETAAPKRTYDMME
                                                                                                                                                                                                                                                                                                                                                                        --EPTPASEATGAPTPPPAPP----SPSAPPPVVPKEEKEEETAAAPPVEEGEEQKP
                                                                                                                                                                                                                                                                                                                                                                                                                                                 PAAEELAVDTGKAEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            --- SGRP-----GRPGERGLPGPPGIKGPAGMPGFPGMKGHRGFDGR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               965 QLKQRAAAIPPIQVTKVHEPPREDAAPTKPAPPAPPPPQNLQPESDAPQQPGSS----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   271 NGEKGETGAPGLK-----GENGLPGDNGAPGPMGPRGAPGERGRPGLPGAAGARGND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          323 GARGSDGQPGPPG------PPGTAG--PPGSPGAKGEVGPAGSPGSNGSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          411 ---ARGPPGPAGTNGIPGTRGPSGEPGKNGAKGEP-GARGERGEAGSPGIPGPKGEDGKD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          508 RGGPGPAGPRGVAGEPGRDG----TPGGPGI------RGMPGSPGGPGN----D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----PPGPQGLQGIPGTGGPPGENGKPGEPGPKGEVGAPGAPGCKGD----SGAPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -----PGTAGIPGARGGAGPP--GPEGGKGPAGPP
                                                                                                                                                                                                                          BEMVEEAEALHASGNEVPRGEC - - - - - SGPATVNNSSDTESIPSPHTEAAKDTGQNGP
                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   365 GORGEPGPGCHAGAQGP--PGPPG-NNGSPGGKGEMGPAGIPGAPGLIG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 548 GKPGPPGS----QGESGRPGPPGPSGP---RGQPGVMGFPGPKGNDGAP-----
                                                                                                                                                 DB 1; Length 1464;
                                                                                                                                                                                                                                                                                                                                                                                                           GGYPGPA-----GPPGPPGPPGSSGHPGSPGSPGYQGPPGEPGQAGPAGPP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HYDROXYLATION (BY SIMILARITY).
HYDROXYLATION (BY SIMILARITY).
HYDROXYLATION (BY SIMILARITY).
HYDROXYLATION (BY SIMILARITY).
INTERCHAIN (BY SIMILARITY).
INTERCHAIN (BY SIMILARITY).
INTERCHAIN (BY SIMILARITY).
INTERCHAIN (BY SIMILARITY).
                                                                                                                                                 Query Match 2.6%; Score 348.5; DB 1; Length Best Local Similarity 20.5%; Pred. No. 0.00014; Matches 359; Conservative 102; Mismatches 578; Indels
                                                                                                                                                                                                                                                                                                 ---PPATLGADGPP-----PGPPTPPRRTSRAPI-----
   859
976
1093
1105
1195
   859
976
1093
1105
1195
1196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ERGP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    217
                                                                                                                                                                                                                          715
   MOD_RES
MOD_RES
MOD_RES
MOD_RES
DISULFID
DISULFID
SEQUENCE
                                                                                                                                                                                                                                                           64
                                                                                                                                                                                                                                                                                                                                                                                                                                                 846
                                                                                                                                                                                                                                                                                                 99/
                                                                                                                                                                                                                                                                                                                                                                          795
                                                                                                                                                                                                                                                                                                                                                                                                             169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            232
   844444
                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ę,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8
```

```
genes."
                                                                                                                                   mouse.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Biol.
                                                                                                                                                                                                                                                                                                                                                                                                                                                     collagen
                                                                                                                                  from
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
1718
                                                                                                                                                     1838
                                                                                                                                                                                                                                                                                                                                                           2118
                                                                                                                                                                                            1893
                                                                                                                                810
                                                                                                                                                                                                                                 1894 KPTVLRSTSTSSPV----RPAATFPPATHCPLGGTLDGVYPTLMEPVLLPKEAPRVARPE 1949
                                                                                                                                                                                                                                                                         1950 RPRADTGHAFLAKPPARSGLEPASSPS----KGSEPRPLVPPVSGHATIARTPAKNLAPH 2005
                                                                                                                                                                                                                                                                                                                                     NGERGPPGPGLPGQPGTAGEPGRDGNPG-----SDGQPGRDGSPGGRGDRGEN 1036
                                                                                                                                                                                                                                                                                                                                                                             1037 GSPGA---PGAPGH-----PGAPGH-----PGPPGPVGPSGKSGDRGETGP---AGPSGAPG 1076
                                                                                                                                                                                                                                                                                                                                                                                                                                        LPAPLYSFPGASCPVLDLRRPPSDLYLPPPDHGAP----ARGSPHSEGGKRSPEPNKTSV 2214
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      KSP------GNTS----QPPAFFSKLTESNSAMVKSKKQEINKKLNTH- 2297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PGPCCGGGAAAIAGVGGEKSGGFSPYYGDDPMDFKINTEEIMSSLKSVNGQIESLISPDG 1251
                                                                                                                                                                                                                                                                                               987
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ---GKDGTSGHPGPIGPPGPRGNRGERGSEG--SPGH------PGQPGPPGPPGA 1191
                                                                                         784
                                                                                                                                                                                                                914
         751
                                                                                                                                                                                                                                                                                                                                                                                                                                                       TVPEHHPHPISPYEHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLDAAAAYYLPRHLAPN
                                                                   PTYPHLYPPYLIRGYPDTAALENROTIINDYITSQQMHHNTATAMAQRADMLRGLSPRES
                                                                                         ----GPAGOPGDKGEGGSP---
                                                                                                          1719 SLALNYAAGPRGIIDLSQVPHLPVLVPPTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLS
                                                                                                                                PGGPTHLTKPTTTSSSERERDRDRERERSILTSTTTVEHAPIWRPGTEQSSGSS
                                                                                                                                                                      811 PGAPGONGEP----GAKGERGAPGEKGEGGPPGPAGPTGSS----GPAGPPGPQGVKGER
                                                                                                                                                                                          GSSGGG-----GSSSRPASHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGIITAVEPS
                                                                                                                                                                                                             ĠŚPĠĠPĠTAGFPĠGRGLPĠPPGNNGNPGPPGPSGAPGKDGPPGPAGNSĠ-----SPĠ
                                                                                                                                                                                                                                                      915 NPGIAGPKGDAGQPGEKGPPGAQGPPGSPGPLG--IAGL--------
                                                                                                                                                                                                                                                                                            952 -----TGARGLAGPPGMPG--PRGSPGPQGIKGESGKP------GASGH
                                                                                                                                                                                                                                                                                                                 HASPDPPAP-----PASASDPHREKTQSKPFSIQELELRSLGYHGSSYSPEGVEPVSPV
                                                                                                                                                                                                                                                                                                                                                                                                 --LLQTAP-----GVKGHQRVVTLAQHISEVITQDYTRHHPQQLSAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LGGGEDGIE----PVSPP-----EGMTEPGHSRSAVYPLLYRDGEQTEPSRMGS
                                                                                                                                                                                                                                                                                                                                                                                                             : ||
PAGARGAPGPQGPRGDKGETGERGSNGIKGHR--------
                                                                                                                                                                                                                                                                                                                                                         SSPSITHDKGLPKHLEELDKSHLEGELRPKQPGPVKLGGEAAHLPHLRPLPESQPSSSP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Craniata; Vertebrata; Euteleostomi;
Sciurognathi; Muridae; Murinae; Mus.
      GPPGASGSPGLQGMPGERGGP---GSPGPKGEKGEPGGAGAD-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (Rel. 01, Created)
(Rel. 17, Last sequence update)
(Rel. 38, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1669 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-FEB-1991 (Rel. 17, Last sequence up
15-JUL-1999 (Rel. 38, Last annotation
Collagen alpha 1(IV) chain precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata;
Mammalia; Eutheria; Rodentia;
NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
                                                                                       762 PAGP-IGPP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          :| | | |
SRKNPARN 1259
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NRNEPEYN 2305
                                                                                                                              --GLPGIAGPRG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CA14 MOUSE
P02463;
21-JUL-1986 (
        713
                                                                                                                                                  1779
                                                                    1659
                           1599
                                                                                                                                                                                                                                                                                                                 2006
                                                                                                                                                                                                                                                                                                                                     886
                                                                                                                                                                                          1839
                                                                                                                                                                                                                                                                                                                                                         2060
                                                                                                                                                                                                                                                                                                                                                                                                2119
                                                                                                                                                                                                                                                                                                                                                                                                                  1077
                                                                                                                                                                                                                                                                                                                                                                                                                                        2159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1192
                                               752
                                                                                                                               785
                                                                                                                                                                                                              863
                                                                                                                                                                                                                                                                                                                                                                                                                                                           1109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COL4AI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT :
CA14_MO'
CA AC
AC DT 210
DT 011
DT 011
DT 02
CC GN CC M
                           ò
                                             엄
                                                               ð
                                                                                  ద
                                                                                                          ò
                                                                                                                           ద
                                                                                                                                          ठ
                                                                                                                                                                  g
                                                                                                                                                                                        à
                                                                                                                                                                                                        පු
                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                        엄
                                                                                                                                                                                                                                                                                                                ઠે
                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                         요
                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                              요
                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
```

```
collagen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SUBUNIT: There are six type IV collagen isoforms, alpha 1(IV)-alpha 6(IV), each of which can form a triple helix structure with 2 other chains to generate type IV collagen network.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     οĘ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Oberbaeumer I., Laurent M., Schwarz U., Sakurai Y., Yamada Y., Vogetli G., Voss T., Siebold B., Glanville R.W., Kuhn K.; Minn C., Siebold B., Glanville R.W., Kuhn K.; Minno acid sequence of the non-collagenous globular domain (NCI) of the alpha I(IV) chain of basement membrane collagen as derived from
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    mouse
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE=88243724; PubMed=3379041;

Killen P.D., Burbelo P., Sakurai Y., Yamada Y.;

Killen P.D., Burbelo P., Sakurai Y., Yamada Y.;

Structure of the amino-terminal portion of the murine alpha 1(IV)

collagen chain and the corresponding region of the gene.";

J. Biol. Chem. 263:8706-8709(1988).

-!- FUNCTION: Type IV collagen is the major structural component of glomerular basement membranes (GBM), forming a 'chicken-wire' meshwork together with laminins, proteoglycans and entactin/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE-86196099; PubMed=3009468;
Sakurai Y., Sullivan M., Yamada Y.;
"Alpha 1 type IV collagen gene evolved differently from fibrillar
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ø
SEQUENCE FROM N.A.
MEDLINE=89197932; PubMed=2703490;
Muthukumaran G., Blumberg B., Kurkinen M.;
Muthe complete primary structure for the alpha 1-chain of mouse collagen IV. Differential evolution of collagen IV domains.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Kurkinen M., Condon M.R., Blumberg B., Barlow D., Quinones S.,
Saus J., Pillajaniemi T.;
"Extensive homology between the carboxyl-terminal peptides of
"Extensive hand alpha 2(IV) collagen.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          type IV collagen genes.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ģ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Nath P., Laurent M., Horn B., Sobel M.E., Zon G., Vogeli G. "Isolation of an alpha 1 type-IV collagen cDNA clone using synthetic oligodeoxynucleotide.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               regulated
                                                                                                                                                                                                                                                                                                                                      Wood L., Theriault N., Vogeli G.; "CDNA clones completing the nucleotide and derived amino sequence of the alpha I chain of basement membrane (type
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Vogeli G.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Burbelo P.D., Martin G.R., Yamada Y.;
"Alpha 1(IV) and alpha 2(IV) collagen genes are
bidirectional promoter and a shared enhancer.";
Proc. Natl. Acad. Sci. U.S.A. 85:9679-9682(1988)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE-89066738; PubMed=3198626;
Kayres P., Wood L., Theriault N., Kurkinen M.,
"Head-to-head arrangement of murine type IV col
J. Biol. Chem. 263:19274-19277(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   bha 1(IV) and alpha 2(IV) collagen."
Biol. Chem. 262:8496-8499(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Chem. 261:6654-6657(1986).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      complementary DNA.";
Eur. J. Biochem. 147:217-224(1985)
                                                                                                                                                                                                                        [2]
SEQUENCE OF 1-1154 FROM N.A.
MEDLINE=88112221; PubMed=3338568;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   [4]
SEQUENCE OF 1276-1669 FROM N.A.
MEDLINE=85127033; PubMed=2578961;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE OF 1441-1669 FROM N.A.
MEDLINE=87250460; PubMed=3597383;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE OF 1149-1424 FROM N.A.
MEDLINE=86301886; PubMed=3755692;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE=89071759; PubMed=3200851;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE OF 1-129 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE OF 1-28 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PARTIAL SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE OF 1-28 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Lett. 227:5-8(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gene 43:301-304(1986).
```

Pred. No. 0.00017;

```
DOMAIN: Alpha chains of type IV collagen have a noncollagenous domain (NC1) at their C-terminus, frequent interruptions of the G-X-Y repeats in the long central triple-helical domain (which may cause flexibility in the triple helix), and a short N-terminal triple-helical 7S domain.

PTM: Prolines at the third position of the tripeptide repeating unit (G-X-Y) are hydroxylated in some or all of the chains.

PTM: Type IV collagens contain numerous cysteine residues which are involved in inter- and intramolecular disulfide bonding. 12 of these, located in the NC1 domain, are conserved in all known type
                 <del>:</del>
```

8 g ò 셤

- - οĘ IV collagens.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@iabb-sib.ch).

셤 8 셤 8 a ò

8

GO; GO:0005604; C:basement membrane; IDA. InterPro; IPR008161; Clg helix. InterPro; IPR00160; Collagen. InterPro; IPR00142; Procollagn4_C. -; NOT ANNOTATED CDS
-; NOT ANNOTATED CDS ANNOTATED_CDS ProDom; PD000007; Clg helix; 6. ProDom; PD003923; ProcellagnC4; 1. SMART; SM00111; C4; 2. AAA51625.1; -. AAA50292.1; -. CAA29946.1; -. EMBL; J03758; AAA37439.1; -. M13043; AAA37346.1; -. J04448; AAA37437.1; -. Pfam; PF01413; C4; 2. Pfam; PF01391; Collagen; 23. AAA37340.1; AAA37342.1; AAA37343.1; AAA37344.1; AAA37345.1; CAA26132.1 Col4al. CGMS4B PIR; A33525; CGN MGD; MGI:88454; M13026; M13027; X06777; X02201; M23333; J04694; M15832; M14042; M12879; M13025; EMBL; EMBL;

ద ò 셤 à 셤

ઠે

DOMAIN) .) (POTENTIAL) (78 COLLAGEN ALPHA 1 (1V) CHAIN.

TRAPLE-HELICAL REGION (NC1).

OR 1548 (BY SIMILARITY).

OR 1551 (BY SIMILARITY).

OR 1652 (BY SIMILARITY).

OR 1665 (BY SIMILARITY).

OR 1665 (BY SIMILARITY).

OR 1665 (BY SIMILARITY).

N-LINKED (GLCNAC. . .) (POTENY A -> P (IN REF. 2).

A -> P (IN REF. 2).

C -> L (IN REF. 2).

C -> L (IN REF. 2).

C -> L (IN REF. 2).

C -> L (IN REF. 2).

C -> L (IN REF. 2).

C -> L (IN REF. 2).

C -> L (IN REF. 2).

C -> H (IN REF. 2).

C -> H (IN REF. 2).

C -> H (IN REF. 2).

C -> H (IN REF. 2).

C -> H (IN REF. 2).

C -> H (IN REF. 2).

C -> H (IN REF. 2).

C -> H (IN REF. 2).

C -> H (IN REF. 2).

C -> H (IN REF. 2).

C -> H (IN REF. 2).

C -> H (IN REF. 2).

C -> H (IN REF. 2).

C -> H (IN REF. 2).

C -> H (IN REF. 2). > Q (IN REF. 2). > H (IN REF. 2). > S (IN REF. 3). 42916B91E52058E9 CRC64; AMINO-TERMINAL PROPEPTIDE Glycoprotein; Collagen; Signal. 160680 Repeat; Hydroxylation; 1669 1669 403 1397 CONFLICT CONFLICT CONFLICT CONFLICT CONFLICT CONFLICT SEQUENCE CARBOHYD CONFLICT CONFLICT DISULFID DISULFID DISULFID DISULFID DISULPID DISULPID DOMAIN SIGNAL PROPEP CHAIN

셤

Š

δ

g

ò

음. 강

셤 8

Length 1669;

DB 1;

2.6%; Score 348;

Query Match

97; SPRGKSRSPAPPADKEAFAAEAQKLPGDP--PCWTSGLPFPVPPREVIKASPHAPDPSAF 1075 1076 SYAPPGHPLPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHV 1135 1136 PYSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEASVL 1195 1196 RGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPSRL 1255 1256 DRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETAAPKRTYDMME 1315 1316 GRVGRAISSASIEGLMGRAIPPERHSPHHLKEQHHIRGSITQGIPRSYVEAQEDYLRREA 1375 1488 HPLDVMADARALERACYEESLKSRPGTASSSG--GSIAR-GAPVIVPELGKP-----RQS 1539 1376 KLLKREGTPPPPPPSRDLTEAYKTQALGPLKLK-----PAHEGLVATVKEAGRSIHE 1427 STVPEHHPHPISPYEHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLDAAAAYYLPRHLAP 1657 207 ----PGPPTPP----RRTSRAPIEPTPASEATGAP 806 PSPSAPPPVV-----PKEEKEETAAAPPVEEGEEQ-----KPPAAEELAVDTG 856 PGPPGPPGEKGOMGSSFQGPKGDKGEQGVSGPPGVPGQAQVKEKGDFAPTGEKGQKGEPG 267 KABEPVKSECTEEAEEGP--AKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSSGAPQD 914 268 FPGVPGYGEKGEPGKQGPRGKPGKDGEKGERGSPGI-----PGDSGYPGLPGRQG-PQG 320 SDSSATCSADEVDEAEGGDKNRLLSPRPSLL--TPTGDPRANASPQKPLDLKQLKQRAAA 972 -----EKGEAG----LPGPPGTVIGTMPLGEKGDRGYPGAP----GLRGEPG- 359 ----RGDIGPPGPPG-----VGPIGPVGEKGQAGFPGGPGSPGLPGPKGEAGKVVPL 644 GTKGTRGPPGAAGYPGNPGLPG1PG0DGPPGPPG1PGCNGTKGERGPLGPPGLPGFSGNP --- PPAP GPPGLPGMKGDPGEILGHVPGTLLKGERGFPGIPGMPGSPGLPGLQGPVGPPGFTGPPGP 973 IPPIOVTKVHEPPREDAAPTKPAPPAPPPONLQPESDAPQOPGS------------KGDPGFPGQ------DGHP PGPQGQKGEPGIG-LPGLKGQPGLPGIPGTPGEKGSIGGPGVPGEQGLTGPPG----------SPHTEAAKDTGQNGPKPP KGDCGGSGCGKCDCHGVKGQKGERGLPGLQGVIGFPGMQGPEGPHGPPGQKGDAGEPGLP 1428 IPREELRHTPELPLAPRPLKEGSITQGTPLKYDTGASTTGSKKHDVRSLIGSPGRTFPPV 645 PGPPGAAGLPGSPGFPGPQGDRGF-PGTPGR----PGIPGEK-----GAVGQPGIGFPGL ------GRPGFNGLPGEIGRPGSP-----GRPGFNGLPGN 1540 PLTYEDHGAPFAGHLP--RGSPVTMREPTPRLQEGSLSSSKASQDRKLTSTPREIAKSPH 405 GPSGRDGAPGPPGP-----PGPPGQPGHTNG-----IVECQPGPPG----Gaps Mismatches 651; Indels 856; -----TGEVGQ-GLPGPKGSPGSI-GLKGERGPP-------GGV--GFPGS-RGECSGPAT-----VNNSSDTESIP-----Best Local Similarity 19.3%; Pro Matches 392; Conservative 133; 441 DOGPPGTPGQPGL----TP-----771 ATLGADGPP-857 360 1018 807 148 28 88 813 208 915 321 543 568 695 1598 셤

셤

ò

g

Extracellular matrix; Connective tissue; Basement membrane;

ઠ 셤 ò 셤

ò

```
15-MAR-2004 (Rel. 43, Last annotation update)
                                                                                                                                                                                                                                                                                                                                              Gerdes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2253
                                                                                                                                                                                         1756 DRLAYLPTAPQPFSSRHSSSPLSPGGPTHLTKPTTTSSSERERDRDRERDREREKSIL 1815
                                                                                                                                                                                                                                                                    1816 TSTTTVEHAPIWRPGTEQSSGSSGSGGG--GSSSRPASHSHAHQHSPISPRTQDALQQ 1873
                                                                                                                                                                                                                                                                                                                                            1874 RPSVLHNTGMKGIITAVEPSKPTVLRSTSTSSPVRPAATFPPATHCPLGGTLDGVYPTLM 1933
                                                                                                                                                                                                                                                                                                                                                                                                                     1934 EPVLLPKEAPRVARPER--PRADTGHAFLAKPPARSGLE------PASSPSKGSE 1980
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1981 PRPLVPPVSGHATIARTPAKNLAPHHASPDPPAPPASASDPHREKTOSKPFSIOELELRS 2040
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2041 LGYHGSSYSPEGVEPVSPVSSPSLTHDKGL-----PKHLEELDKSHLEGELRPKQPGPV 2094
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----GLPGIGIPGRPGDKGDQGLAGFPGSPGEKGSKGSAGTPG--MPGSPGPR 1099
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2095 KLGGEAAHLPHLRPLP-ESQPSSSPLLQTAPGVKGHQRVVTLAQHISEVITQDYTRHHPQ 2153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              134 --EAGLPGT----PGPTGPAGOKGEPGSD-GIPGSAGEKGEOGVPGKGFPGFPGSK-GDK 1185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1186 GSKGEVGFPGLAGSPGIPGVKGEQGF--MGPPGPQGQPGLPGTPGHPVEGPKGDRGPQGQ 1243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PSRMGSKSPGNTSQPPAFFSKLTESNSAMVKSKKQEIN-KKLNTHNRNEPEYNISQPGTE 2312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1244 PGLPGHPGP---MGPPGF-------PGINGPKGDKGNQGWP----GAPGVP 1280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----IFNMPAITGTGLMTYRSQAVQEHASTNMGLEAII----RKALMG----KYDQW 2357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----GSKGDMGLPGVPGFQGQKGLPGLQGVKGDQG 1332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EESPPLSANAFNPLNASASLPAAMPITAADGRSDHTLTSPGGGGKAKVSGRPSSRKAKSP 2417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ------VPGPKGLQGPPGPPGPYDVIKG 1357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2418 APGLASGDRPPSVSSVHSEGDCNRRTPLTNRV------WEDRPSSAGST-PF-PYN 2465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1358 EPGLPGPEGPPGLKGLQGPPGPKGQQGVTGSVGLPGPPGVPGFDGAPGQKGETGPFGPPG 1417
                                   -----RGYPDTAALENRQTIINDYITSQQMHH 1697
                                                                                                                1698 NTATAMAQRADMLRGLSPRESSLALNYAAGPRGIIDLSQVPHLPVL--VPPTPGTPATAM
                                                                                                                                                   ------MPGPKGDKGSQGLPGLTGQSGL-----PGLPGQGTPGVPGPPG-SK
                                                                                                                                                                                                                                                                                                       -----BHG----EMKGDRGDVGLPGMPGSMEHVDMG---SMKGQKGDQG
                                                                                                                                                                                                                                                                                                                                                                                  955 EKGQIGPTGDKG------SRGDPGTPG-------955
                                                                                                                                                                                                                                                                                                                                                                                                                                                         ----vpgkdggaghpgpgpkgbpg---bsgrpgspglpgpkgsvggmglpgspgekg--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1100 GSPGNIGH-PGSPGLPGEKGDKGLPGLDGVPGVKG--------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QLSAPLPAPLYSFPGASCPVLDLRRPPSDLYLP----PPDHGAPARGSPHSEGGKRSPE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PNKTSV----LGG-----GEDGIEPVSPPEGMTEPGHSRSAVYPLLYRDGE---QTE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PLIMRLQAGVMASPPPPGLPAGSGPLAGP------HHAWDEEPKPLLC 2507
 ----LOGIRG------DPG--PPGVOGPAGP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ----VPGIPGSQGVPGSPGEKGA------KGEKGQS
                                                                        - PGVPGIGPPGAMGPPGGEGPPGSSGPPGIKGEKGFPGFPGLD-
                                                                                                                                                                                                                              879 GEMGVMGTPGOPGSPGPAGTPGLPG-----
                                     1658 NPTYPHLYPPYLI------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GPKGDPGFQGMPGIGGSPGIT-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1333 DQG------
                                                                                                                                                                                                                                                                                                         907
                                                                                                                                                                                                                                                                                                                                                                                                                                                           976
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1027
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2313
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2466
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1053
                                                                          964
                                                                                                                                                     838
                                                                                                                                                                                                                      8 & 8
                                                                                                                                                                                                                                                                                                                               8 8
                                                                                                                                                                                                                                                                                                                                                                                                      8 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          6 6 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ठे
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8 B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                    g
                                                                                                                                         셤
                                                                                                                                                                             ∂
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8
                                   8
                                                                                                              ઠે
```

Z

STANDARD;

KI67_HUMAN ID_KI67_HUMAN AC P46013; DT 01-NOV-1995 (DT 01-NOV-1995 (

RESULT 58

(Rel. 32, Created) (Rel. 32, Last sequence update)

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation between the Swinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             coved. Usage by and for commercial
(See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                               -!- FUNCTION: Thought to be required for maintaining cell proliferation.
-!- SUBCELLULAR LOCATION: Nuclear. Predominantly localized in the G1 phase in the perinucleolar region, in the later phases it is also detected throughout the nuclear interior, being predominantly localized in the nuclear matrix. In mitosis, it is present on all
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Isocial P46013-2; Sequence-VSP 004298;
Isocial STAGE: Expression of this antigen occurs
DEVELOPMENTAL STAGE: Expression of this antigen occurs
preferentially during late 61, S, G2 and M phases of the cell
cycle, while in cells in G0 phase the antigen cannot be detected.
SIMILARITY: Contains 1 FHA domain.
                                                                                                                                                                          "The cell proliferation-associated antigen of antibody Ki-67: a verlarge, ubiquitous nuclear protein with numerous repeated elements, representing a new Kind of cell cycle-maintaining proteins."; J. Cell Biol. 123:513-522(1993).
                                                      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            16 X 122 AA APPROXIMATE REPEATS
                                                                                                                                                C., Becker M.H.G., Key
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PROSITE; PS50006; FHA DOMAIN; 1.
Cell cycle; Antigen; Nuclear protein; ATP-binding; Repeat;
                                                                                                                                                                                                                                                                                    Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                         chromosomes.
ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GO; GO:0008283; P:cell proliferation; TAS. GO; GO:0000074; P:regulation of cell cycle; TAS. InterPro; IPR000253; FHA. InterPro; IPR008984; SMAD_FHA.
                                                                                                                                                                                                                                                                                                                                                                                                                                   Name=Long;
IsoId=P46013-1; Sequence=Displayed;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            or send an email to license@isb-sib.ch)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                entities requires a license agreement
                                                                                                                  SEQUENCE FROM N.A.
MEDLINE=94043435; PubMed=8227122;
Schlueter C., Duchrow M., Wohlenberg
Flad H.-D., Gerdes J.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Polymorphism.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; X65550; CAA46519.1; -.
EMBL; X6551; CAA46520.1; -.
EMBL; X94762; CAA64388.1; -.
                                                                                                                                                                                                                                                     SEQUENCE OF 1-31 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PIR; A4866; A4866.
Genew; HGNC:7107; MK167.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Alternative splicing; F DOMAIN 27 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               11112
1234
1356
1477
1598
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1842
1964
2086
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam; PF00498; FHA; 1.
SMART; SM00240; FHA; 1
                                             Homo sapiens (Human)
                                                                                      NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1000
1122
1124
1244
1366
1608
1730
1974
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Name=Short;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3K; P46013; -.
               Antigen KI-67.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  76741;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REPEAT
```

	1299 PEKEVGLEKDIII FVOTPVOKLDLTENLIGSKRR	
8 8 8 8 8 8 8 8	4 6 4 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	6 B 6 B 6 B 6 B 6 B 6
FT REPEAT 2096 2204 10. FT REPEAT 2214 2326 11. FT REPEAT 2214 2326 11. FT REPEAT 2219 2869 13. FT REPEAT 2699 2808 15. FT REPEAT 2699 2808 15. FT REPEAT 2699 2808 15. FT REPEAT 2699 2808 16. FT NP BIND 3034 APT POTENTIAL). FT VARIANT 3150 3150 T -> S (in dbSNP: 11106). FT VARIANT 3150 3150 T -> S (in dbSNP: 11106). FT FT VARIANT 3217 3217 K -> B (in dbSNP: 8473). FT VARIANT 3217 3217 K -> B (in dbSNP: 8473). FT COLETY Match 2.6%; Scote 348; DP 14859. Query Match 2.6%; Scote 348; DP 1; Length 3256; Best Local Similarity 18.7%; Pred. NO. 0.00032; Matches 537; Conservative 352; Mismatches 1129; Indels 852; Gaps 133;	325 SVQTPSKAVGASFPLYEPAKWKTPVQYSQQONSPQKHKNKDLYTTGRRESVNLGKS 52HLSPGSIIQPQRRRPSLLSEPQPGNERSQEHHLRPESHSYLPEL	

2, Last annotation update) condensation inducer in the nucleus (Acinus)

42,

10-OCT-2003 (Rel. 42 Apoptotic chromatin

```
1737
                                                                                                                      2085
                                                                                                                                                                                                                                         2176
                                                                                                                                                                                                                                                                                                      2218
                                                                                                                                                                                                                                                                                                                                                                                                                                                      2043
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2103
1952 EDLAGFKELFQTPGHTEESMTDDKITEVSCKSPQ---PDPVKTPTSSKQRLKISLGKVGV 2008
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2417
                                                                                                                                                                                                                                                                         1904
                                                                                                                                                                                                                                                                                                                                  ------VLLPKEAPRVARP- 1948
                                                                                                                                                                                                                                                                                                                                                                                         ERPRADIG-----PASTAKPPARSGLEPASSPSKGSEPRPLVP-----PVSGHATIA 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PHLRPLPESQPSSSPLLQTAPGVK-----GHQRVVTLAQHISEVITQD-----YTRH 2150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PGSKROPOTPKEKAEALEDLVGFKELFQTPGHTEESMTDDKITEVSCKSPQPESFKTSRS 2477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2151 HPQQLSAPLPAPLYSFPGASCPVLDLRRPP---SDLYLPPPD----HGAPARGSPHSEGG 2203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   QTEPSRMGSKSPGNTSQPPAFFSKLTESNSAMVKSKKQEINKKLNTHNRNEPEYNISQPG 2310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      --TUMGLEAIIRKALMGKYDQWE-ESPPLSANAFNPLNASASL-------P 2378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GHTQESLTA-----GKATKIPCESPPLEV----VDTTASTKRHLRTRVQKVQVKEEP 2737
                                                                                                        -----SKRQPRTPKGKAQPLEDL--AGLKELFQTPVCTDKPTTHEKT-TK
                                                                                                                                                                                                                                                                                                                                                    2219 IACRSPQPDPVGT------PTIFKPQSKRSLRKADVEESSLALRKRTPSVGKAM
                                                                                                                                                                                                                                                                                                                                                                                                            YRSHIPLA-FDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAALENR
                                                                                     QTIINDYITSQQMHHNTATAMAQRADMLRGL----SPRESSLALNYAAGPRGIIDLSQV
                                                                                                                                                 P-HLPVLVPPTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTHLTKPTTTSSSER
                                                                                                                                                                              2086 PDHT-----BESTIDDKTIKIACKSPP------PESMDTPTST---R
                                                                                                                                                                                                          1797 ERDRDRERBREREKSI---LISTITVEHAP-----IWRPGTEQSSGSSGSSGGG
                                                                                                                                                                                                                                        2119 RRPKTPLGKRDIVEELSALKQLTQTTHTDKVPGDEDKGINVFRETAKQKLDPAASVTG--
                                                                                                                                                                                                                                                                    1846 GSSSRPASHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGII-TAVEPSKPTVLRSTSTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                    RTPAKN------LAPHHASPDPPAPASASDPHREKTQSKPFSIQE--LELRSLGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2324 OTPGTDKPTTDEKTTKIACKSPQPDPVDTPASTK--QRPKRNLRKADVEEEFLALR----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2478 SKQRLKIPLVK------VDMKEBPLAVSKLTRTSGETTQTHTEPTGDSKSIKAF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         KRSPE---PNKTSVLGGGED-----GIEP-VSPPEGMTEPGHSRSAVYPLLYRDGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2526 KESPKQILDPAASVTGSRRQLRTRKEKARALEDLVDFKELFSAPGHTEESM-----T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2044 HGSSYSPEGVEPVSPVSSPSLTHDKGLPKHLEBLDKSHLEGELRPKQPGPVKLGGEAAHL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAMPITAADGRSDHTLTSPGGGGKAKVSGRPSSRKAKSPAPGLASGDRPP 2428
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 --GKTTOTH-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TEIFNMPAITGTGLMTYRSQA-----VQEHAS-----
                                                                                                                                                                                                                                                                                                                               1905 SPVRPAATFPPATHCPLGGTLDGVYPTLMEP-
                                                           2009 KEEVLPVGKLTOTS-----
                                                                                                                                                 1738
                                                                                                                                                                                                                                                                                                                                                                                         1949
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2104
                                                                                                                                                                                                                                                                                                                                                                                                                                                   1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2251
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2379
                                                                                                                                                                                                                                                                                                   2177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2630
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2338
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2738
셤
                             ò
                                                        셤
                                                                                     à
                                                                                                             셤
                                                                                                                                               ò
                                                                                                                                                                        셤
                                                                                                                                                                                               õ
                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                  ઠે
                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                            ద
                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à. a
```

ACIN_MOUSE STANDARD; PRT; 1338 AA. 09JIX8, 09CSN1, 09CSR3, 09CSR7; 09R046; 09R047; 28-FEB-2003 (Rel. 41, Created) 28-FEB-2003 (Rel. 41, Last sequence update)

RESULT 59
ACIN MOUSE
ID ACIN MC
AC Q9JIX8;
DT 28-FEBDT 28-FEB-

```
REQUENCE OF 1-1190 FROM N.A. (ISOPORNS 1; 2 AND 4).

SEQUENCE OF 1-1190 FROM N.A. (ISOPORNS 1; 2 AND 4).

RA SEQUENCE-234663; PubMed=1246681;

RA MEDLINE-2534663; PubMed=1246681;

RA MEDLINE-2534663; PubMed=1246681;

RA MEDLINE-2534663; PubMed=1246681;

RA MEDLINE-2534663; PubMed=1246681;

RA MEDLINE-254663; PubMed=1246681;

RA Baldarelli R., Hall D.P., Bult C., Hume D.A., Canonbach C., Gojobori T., Raldarelli R., Hill D.P., Bult C., Hume D.A., Canonbach C., Gojobori T., Baldarelli R., Fanapin A., Matsuda H., Batalov S., Beisel K.W., Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W., RA Baldarelli R., Bragani T.A., Fletcher C.F., Forrest A., Frazer K.S., Gasterland T., Gariboldi M., Gissi C., Godzik A., Gough J., Gasterland T., Gariboldi M., Gissi C., Godzik A., Gough J., Ronagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A., Kawaji H., Kawaswa Y., Kedzierski R.M., King B.L., Kanaji H., Kawaswa K., Marchionni L., McKenzie L., Miki H., Rapashina T., Numata K., Oxido T., Paran W.J., Pertea G., Pesole G., Petrovsky N., Pillai R., Pontius J.U., Oi D., Ramachandran S., Sultana R., Takenaka Y., Taylor M.S., Tesadale R.D., Tomita M., Sandelin A., Schneider C., Semple C.A., Secou M., Shimada K., Sultana R., Takenaka T., Wahlestedt C.A., Warden W., Shimada K., Avan K., Marki K., Kawai J., Yang I., Yang I., Yang I., Yang I., Yang I., Yang I., Yang I., Yang I., Yang I., Yang I., Anakawa T., Konno H., Nakamura M., Sakazume N., Sakazume N., Sakazume N., Sakazume N., Sakazume N., Sakazume N., Sakazume N., Sakazume N., Sakazu J., Alazawa K., Arakawa T., Rohino M., Waterston R., Lander R., Shinagawa I., Rhinayazaki A., Sakai K., Sasaki D., Shinagawa G., Rhinayawa I., Rhinayasi of the mouse transcriptome based on functional annotation of R. Mallysis of the mouse transcriptome was every condense of functional annotation of R. Thurs 420.561-561-573 (2002)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     to frameshifts
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IsoId=09JIX8-4; Sequence=VSP_004032; PTW: Undergoes proteolytic cleavage; the processed form is active, contrary to the uncleaved form (By similarity). SIMILARITY: Contains 1 SAP domain.
                                                                                                                                                                                                                                                                                                                                                                      chromatin
                                                                           Bukaryota; Metazoas;
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Butheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
                                                                                                                                                                                                                   Tsujimoto |
apoptotic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -!- FUNCTION: Induces apoptotic chromatin condensation after activation by CASP3 (By similarity).
-!- SUBCELUMAR LOCATION: Nuclear (By similarity).
-!- ALTERNATIVE PRODUCTS:
                                                                                                                                                                                                                                                                                                                                            Mamoru A., Setsuko S., Yoshihide T.; Manlecular cloning of murine acinusL, a gene for apoptotic
                                                                                                                                                                                                                                                                                                                                                                                                          Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     due
                                                                                                                                                                                                          Sahara S., Aoto M., Eguchi Y., Imamoto N., Yoneda Y., "Acinus is a caspase-3-activated protein required for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence=VSP_004030, VSP_004033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CAUTION: Ref.2 sequence differs from that shown in position 110 and 112.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Event=Alternative splicing; Named isoforms=4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IsoId=Q9JIX8-3; Sequence=VSP_004031;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence=Displayed;
                                                                                                                                                                         2 AND 3).
                                                                                                                                                                                           MEDLINE=99418558; PubMed=10490026;
                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A. (ISOFORM 1)
                                                                                                                                                                         (ISOFORMS
                                                                                                                                                                                                                                                         chromatin condensation."
                                                                                                                                                                                                                                                                                Nature 401:168-173(1999)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Name=1; Synonyms=L;
Isold=Q9JIX8-1; S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Name=2; Synonyms=S;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Synonyms=S'
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Isold=09JIX8-2;
                                                              Mus musculus (Mouse)
                                                                                                                                                                     FROM N.A.
                                                                                                                          NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                           condensation."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Name=4
                                                                                                                                                                     SEQUENCE
```

Db 204PRSSSFSEEKGESDDEK	Qy 522 EKEKEKEAERKEEERPEVENDKEDLLKEKTDDTSGEDNDEKEAVASKGRKTANSGGRR 578 Db 276 EBEEDDDDEBEBEVDEAQKSREARAPTLKQ-FEDEEGEERTRAKPEKVVDEKPLNIRSQE 334 Qy 579 KGRITTSMANNEANSEEAITPQQSAELASMELNESSRWTEEEMETAKKGLLEHGR 632 Db 335 KGELEKGRYTRSO-ERARRSHLARQOORKETOVSLPDEBENY 377	633 NWSAIARWVGSKTVSQCKNFYFNYKKRQNLDEILQQHKLKWEKERNARRKKKKAPAAASE	753 SPHTEAAKDTGQ-NGPKPPATLGADGPPGPPTPPRRTSRAPIEPTPASEA	858 AEEPVKSECTEBAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSSGAPQDSDS		038	1189 AQEASURGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYKGTITRIIG 1189 AQEASURGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYKGTITRIIG 1249 EDSPSRLDRGREDSLPKGHVISVEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHE 1249 EDSPCRLDRGREDSLPKGHVISVEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHE 1249 EDSPSRLDRGREDSLPKGHVISVEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHE 1249 EDSPTENDGREDSLPKGHVISVEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHE 1304 TAAPKBTYDMMEGRVGRAISSASIEGLMGRAIPPERHSPHHLKEQHHIRGSITQGIPRSY 1304 TAAPKBTYDMMEGRVGRAISSASIEGLMGRAIPPERHSPHHLKEQHIRGSITQGIPRSY 1304 TAAPKBTYDMMEGRVGRAISSASIEGLMGRAIPPERHSPHHLKEQHIRGSITQGIPRSY 1364 VEAQEDYLRREAKLLKREGTPP
between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).	EMBL; AF124725; AAD56723.1; EMBL; AF124729; AAD56727.1; EMBL; AF168782; AAF89661.1; ALT FRAME. EMBL; AK011698; -; NOT ANNOTATED_CDS. EMBL; AK012099; BAD28030.1; EMBL; AK012037; BAD28171.2; EMBL; AK050467; BAD28171.2;	MGI:18918 0:000573C 0:0016887 0:003676 0:0030263 0:0030263	erpro; IPR003034; SAP. n; PR02037; SAP; 1. RT; SM00513; SAP; 1. ptosis; PS50800; SAP; 1. ptosis; Nuclear protein; Alternative splicing. AIN 72 106 SAP. AIN 142 443 GLU-RICH. AIN 569 667 SER-RICH.	1113 1130 1131 1338 1093 1094 IC 1 757	164 204 758 766	CONFLICT 244 244 T -> A (IN REF. 2). CONFLICT 515 515 T -> A (IN REF. 2). CONFLICT 536 536 F -> L (IN REF. 2). CONFLICT 569 568 H -> Y (IN REF. 2). CONFLICT 599 599 V -> A (IN REF. 2). CONFLICT 729 729 S -> I (IN REF. 2). CONFLICT 729 729 S -> I (IN REF. 2). CONFLICT 773 773 T -> A (IN REF. 2). CONFLICT 829 829 MISSING (IN REF. 1; AAD56723). CONFLICT 829 836 Q -> R (IN REF. 3; BAB28030). CONFLICT 1035 1035 G -> R (IN REF. 3; BAB28030).	1338 AA; 150690 MW; B912D9CBSTSOFBI 2.64; Score 346; DB 1; 9; Conservative 206; Mismatches 570; POVELPLYNQPSDTRQYHENIKINQAMR-KKL] :: : : : PRAMPEAVCTDPSTSRKMABLEEVTLDGKPLQALRYTDLA DQLMEALEK-KVERIENNPRRAKESKVREYYEKQFI KGALMLENLQKHSTPHAAFQPNSQIGEENSQNSFIKQY- CLSMSAARSEHBVSEILDGLSBQRLEKQMRQI :: : : : :

6 6 6 6 6 6

us-09-522-753-5.rsp

÷

```
1189
                                                                                                                                                                     1654
987 GVSITIDDPVRTAQVPSPPRGKISNIVHISNLVRPFTLGQLKELLGRTGTLVEEAFWIDK 1046
                      1459 -----YDT--GASTTGSKKHDVRSLIGSPGRTFPPVHPLDVMADARALERACYEESL 1508
                                                                          1567
                                                                                                                       --EIAK 1594
                                                                                                                                                                                            1190 SKEKKSEKKEKAQEEPPAKLLD-----DLFRK------TKAAPCIYWLP-- 1227
                                                                                                                                                                                                                     1714
                                                                                                                                                                                                                                         1715 PRESSLALNYAAGPRGIIDLSQVPHLPVLVPPTPGTPATAMDRLAYLPTAPQPFSSRHSS 1774
                                                                                                                                                                                                                                                                                          ------EQKER-----EKEAERERN 1271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Oshima M., Sugiyama H., Kitagawa K., Taketo M.;
"APC gene meseenger KNA: novel isoforms that lack exon 7.";
Cancer Res. 53:5589-5591(1993).
-!- FUNCTION: Tumor suppressor. Promotes rapid degradation of CTINNB1
and participates in Wnt signaling. APC activity is correlated with
                                                                                       1130 PRAEQREÓERAVREQWAERERERERERTRSEREWDRDKVREGPRSRSRSRDRRRKERAK
                                              1047 IKSHCFVTYSTVEEAVATRTALHGVK------WPQSNPKFLCADYAEQDELDYH---
                                                                      1509 KSRPGTASSSGGSIARGAPVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSPVTMREPTP-
                                                                                                                                                                    SPHSTVPEHHPHPISPYEHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLDAAAAYYLPRH
                                                                                                                                                                                                                 LAPNPTYPHLYPPYLIRGYPDTAALENRQTIINDYITSQQMHHNTATAMAQRADMLRGLS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Su L.-K., Kinzler K.W., Vogelstein B., Preisinger A.C., Moser A.R., Luongo C., Gould K.A., Dove W.F.; Science 256:1114-1114(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND VARIANTS.
STRAIN=C57BL/6J, and CAS7/EI; TISSUE=Brain;
MEDLINE=525631011, Pubmed=13150108;
Su L.-K., Kinzler K.W., Vogelstein B., Preisinger A.C., Moser A.R.,
Luongo C., Gould K.A., Dove W.F.;
"Multiple intestinal neoplasia caused by a mutation in the murine
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia, Butheria, Rodentia, Sciurognathi, Muridae, Murinae, Mus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE OF 1-45 FROM N.A.
STRAIN-BALB/c, TISSUE-Liver;
Dicker F., Lambertz S., Reitmair A., Ballhausen W.G.;
"The murine APC gene: alternative aplicing of 5' untranslated
                                                                                                                                                                                                                                                                                                                                                                                                                                    16-OCT-2001 (Rel. 40, Last sequence update)
10-OCT-2003 (Rel. 42, Last annotation update)
Adenomatous polyposis coli protein (APC protein) (mAPC).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Submitted (OCT-1993) to the EMBL/GenBank/DDBJ databases.
                                                                                                                       --KLTSTPR-
                                                                                                                                                                                                                                                                                                                 1775 SPLSPGGPTHLTKPTTTSSSERERDRDRERDR-DREREK 1812
                                                                                                                                                                                                                                                                                                                                       1272 RQLE-----REKRREHSRÉRERDRERERDRERER 1303
                                                                                                                                                                                                                                                                                                                                                                                                    2845 AA
                                                                                                                      1568 ----RLQEGSLSSSKASQDR-----
                                                                                                                                                                                                                                                                                                                                                                                                   PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE=94061824; PubMed=8242607;
                                                                                                                                                                                                                                                                                                                                                                                                          061315; 062044;
16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last seq
10-OCT-2003 (Rel. 42, Last ann
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       gene.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   homolog of the APC gene."
Science 256:668-670(1992)
                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ALTERNATIVE SPLICING.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Mus musculus (Mouse)
                                                                                                                                                                                                                                                                                          EREEE----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 region segments
                                                                                                                                                                                                                                                                                                                                                                                                    APC MOUSE
                                                                                                                                                                                                                   1655
                                                                                              1095
                                                                                                                                                                    1595
                                                                                                                                                                                                                                                                                         1253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ERRATUM.
                      ò
                                             셤
                                                                     ò
                                                                                          ద
                                                                                                                    8
                                                                                                                                          吕
                                                                                                                                                                    ò
                                                                                                                                                                                       g
                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                     g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                  IsoId=Q61315-4; Sequence=VSP_004116, VSP_004117; TISSUE SPECIFICITY: Expressed in liver, spleen, kidney, heart, lung, brain, stomach, intestine, testis and ovary. PTM: Phosphorylated by GSK3B (By similarity). SIMILARITY: Contains 7 ARM repeats.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Missing (in isoform 2 and isoform 4). /FTId=VSP_004116.
its phosphorylation state (By similarity).
SUBUNIT: Forms homooligomers. Associates with catenins. Binds
axin (By similarity).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (in isoform 3 and isoform 4).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GO; GO:0005737; C:cytoplasm; IDA.
GO; GO:0005634; C:nucleus; IDA.
GO; GO:0005634; C:nucleus; IDA.
GO; GO:000913; P:beta-catenin binding; IDA.
GO; GO:0009952; P:anterior/posterior pattern formation; IMP.
GO; GO:0009953; P:axis specification; IMP.
GO; GO:0009953; P:dorsal/ventral pattern formation; IMP.
GO; GO:0016055; P:WIT receptor signaling pathway; IDA.
InterPro; IPR000225; Armadillo.
Pfam; PF00014; Armadillo_seg; 4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       V -> I (IN STRAIN CAST/EI).
V -> I (IN STRAIN CAST/EI).
Y -> F (IN STRAIN CAST/EI).
A -> T (IN STRAIN CAST/EI).
G -> A (IN STRAIN CAST/EI).
H -> Q (IN STRAIN CAST/EI).
T -> A (IN STRAIN CAST/EI).
T -> A (IN STRAIN CAST/EI).
T -> A (IN STRAIN CAST/EI).
T -> A (IN STRAIN CAST/EI).
W, 145CA73CF570A499 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SMART; SM00185; Akw; J.

PROSITE; PS50176; ARM REPEAT; 1.

Wht signaling pathway; Anti-oncogene; Phosphorylation; Alternative splicing; Repeat; Coiled coil.

DOMAIN

1 61

COILED COIL (POTEWIAL).

245

COILED COIL (POTEWIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ARM 2.
ARM 3.
ARM 4.
ARM 4.
ARM 5.
ARM 6.
ARM 7.
ARM 7.
ARP 7.
ARP/GIJU-RICH (ACIDIC).
ASP/GIJU-RICH (ACIDIC).
                                                                 ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             004117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HIGHLY CHARGED.
                                                                                                                                                                                IsoId=Q61315-2; Sequence=VSP_004116;
                                                                                                                                                                                                                          IsoId=061315-3; Sequence=VSP_004117;
                                                                                                                                     IsoId=Q61315-1; Sequence=Displayed;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LEU-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Missing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ARM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL, M88127, AAB59632.1, -.
EMBL, U02937, AAA03443.1, -.
PIR, I49505, 149505.
HSSP, Q02248, 38CT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2834
1156
1575
1891
276
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1330
1618
2294
2496
2523
2813
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MGD; MGI:88039; Apc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1330
1618
2294
2496
2523
2813
2845
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          556
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         243
                                                                                                                                                                                                                                                 Name=4;
                                                                                                                                                                                                     Name=3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VARIANT
VARIANT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VARIANT
VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                 1
                       <u>+</u>
```

AA; 311086 MW;

ò	901 GRATTAKSSGAPQDSDS	GRATTAKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRANASPQKP 960
qq	1499FSC	:: : -FSCSSSLSALSLDEPFIQKD 1518
ò	961 LDLKQLKQRAAAIPPIQVT	VTKVHEPPREDAAPTKPAPPAPPPPQNLQPESDAPQQP 1015
qq	: : 1519 VELRIMPPVC	
ò	1016GSSPRGKSRE	GSSPRGKSRSPAPPADKEAFAAEAQKLPGDPPCWTSGLPFPVPPREVIK 1064
q ₀	1572 ILEECIISAMPTKSSRK	(AKKLAQTASKLPPPVARKPSQLPVYKLLPAQNRLQ 1623
ò	1065 ASPHAPDPSAFSYAPPC	ASPHAPDPSAFSYAPPGHPLPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLERQI 1122
QQ	1624 AQKHVSFTPGDD	DDVPRVYCVEGTPINFSTATSLSDLTIESPP 1664
δ	1123 GAISQGMSVQLHVPYSE	GAISQGMSVQLHVPYSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPE 1182
Dp	1665 NELATGDGVRAGIQSGEFEKRDTIPT	FEKRDTIPTEGRSTDDAQRGKIS 1704
ò	1183 SLGVPTAQEASV	TAQEASVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHG 1232
Dp	1705 SIVTPDLDDNKAEEGDI	IVTPDLDDNKAEEGDILAECINSAMPKGKSHKPFRVKKIMDQVQQASSTSSGANKNQ 1762
ò	1233TPADVLYR	TPADVLYKGTITRI-IGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEG 1282
q ₀	1763 ИОТККККРТЅРVКРМРС	VDTKKKKPTSPVKPMPQNTEYTRVRKNTDSKVNVNTEETF 1803
ò	1283 GMSVTQCSKEDGRSSSC	GMSVTQCSKEDGRSSSGPPHETAAPKRTYDMMEGRVGRAISSASIEGLMGRAIPPERHSP 1342
qq	1804SDNKDSKK	SDNKDSKKPSLQTNA-KAFNEKLPNNEDRVRGTFALDSP 1841
ò	1343 HHLKEQHHIRGSITQGIPRSYVEAQEDYL	RREA
QO	1842 HHYTPIEGI	
È	1390 SRDLTEAYKTQALGPLK :	SRDLTBAYKTQALGPLKLKPAHEGLVATVKEAGRSIHBIPREBLRHTPELPLAPRPL 1446
q	1885 SKD-SEAKVICRPEPNSSQQAASKSQASIKH	SQQAASKSQASIKHPANRAQSKPVLQKQPTFPQSS 1935
ò	1447 KEGSITQGTPLKYDTGA :	KEGSITQGTPLKYDTGASTTGSKKHDVRSLIGSPGRTFPPVHPLDVMADARALERACYEE 1506
qq	1936 KDGPDRGAAT	ATDEKLQNLAIENTP-VCFSRNSSLSSLSDIDQ-ENNNNKE 1982
ò	1507 SLKSRPGTASSSGGSIA	SLKSRPGTASSSGGSIARGAPVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSPVTMREPT 1566
qq	1983 SEPIKEAEPANSQG	EPSKPQASGYAPKSFHVEDT 2016
È	1567 PRLQEGSLSS-SKASQDRKL	AKSPHSTVPEHHPHP
qa	2017 PVCFSRNSSLSSLSIDS	PVCFSRNSSLSSLSIDSEDDLLQECISSAMPKKKRPSRLKSESEKQSPRKV 2067
ò	1619 SGVDLYR	DLYRSHIPLAFDP151PRGIPLDAAAAYYL 1651
q	2068 GGILAEDLTLDLKDLQR	GĠILAEDLTLDLKĎĹQŘPDSEHÁŘSŘGSENFDWKAIQEGANSÍVSSLHQAÁÁÁAACLSRQ 2127
'n	1652	PRHLAPNPTYPHLYPPYLIRGYPDTAALENROTIIND 1688
qa	2128 ASSDSDSILSLKSGISI	ASSDSDSILSLKSGISLGSPFHLTPDQEEKPFTSNKGPRILKPGEKSTLEAKKIESENKG 2187
ò	1689 YITSQQMHHNTATAMAC	YITSQOMHHWTATAMAQRADMLRGLSPRESSLALNYAAGPRGIIDLSQVPHLPVLVPPTP 1748
qq	2188 IKGGKKVYKSLITGKIRSNSEIS-	SNSEISSQMKQPLPTNMPSISRGRTMIHIPGL 2236
ò	1749 GTPATAMDRLAYLPTAF	GTPATAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTHLTKPTTTSSSERERDRDRERDRDR 1808
qa qa	. 2237	RNSSSSTSPVSKKGPP-LKTPASKSPSEGPGATTSPRG-TK 2275
δ	1809 EREKSILTSTTVEHAP	EREKSILISTITVEHAPIWRPGTEQSSGSSGSSGSSGSSSRPASHAHQHSPISPRIQ 1868
qq	2276 PAGKSELS	PITRÓTSQISGSNKGSSRSGSKDSTPSRPTQ-OPLSRPMQ 2322
à.	1869 DALQQRPSVLHNTGMKG	DALQQRPSVLHNTGMKGIITAVEPSKPTVLRSTSTSSPVRPAATFPPATHCPLGGTLDGV 1928
	•	

```
8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     음
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2089
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2642
S--PGRNSI--SPGRNGI---SPPNKLSQLPRTSSPSTA-----STKSSGSGKMSYT 2367
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2723
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2208 BPNKTSVLGGGEDGIEPVSPPEGMTEPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQ 2267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TAADGRSDHTLTSPGGGGKAKVSG---RPSSRKAKSPAPGLASGDRPPSVSSVHSEGDCN 2440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SSSS-SKH--SSPSGTVAARVIPFNYNPSPRKSSADSTSARPSQIPTPVSTNTKKRD-- 2819
                                                                                                                                                                                                                                                                                                                                                            1975 PSKGSE----PRPLVPPVSGHATIARTPA-KNLAPHASPDPPAPPASASDPHREKTQSK
                                                                                                                                                                                                                                                                                                                                                                                                                                2423 KSSGSESDSSERPALVRQSTFIKEAPSPTLRRKLEESASFESLSPSSRPDSPTRSQAQT-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2540 ---PINRAGTWKREHSKH-----SSSLPRVSTWRRTGSSSSILSASSESEKAKSED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | : | : | : : | | : : | | : : | : | : : | | : : | | : : | | : : | : | : | : : | | : : | : | : | : | : : | : | : | : | : 
                                                                                                                                                   1929 YP-----TLMEPVLLPKEAPRVARPERPRADTGHAFLAKPPARSGLEP----ASS
                                                                                                                                                                                                                                                       2368 SPGRQLSQONLTKOASLSKNASSI-----PRSESASKGLNQMSNGNGSNKKVELSRMSST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2030 PFSIQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDKGLPKHLEELDKSHLEGELRPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2482 PVLSPSLPDMSLSTH-PSVQAGGWRKLPPNLSPTIEYNDGRPTKRHDIARSHSESPSRL-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QPGPVKLGG--EAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQRVVTLAQHISEVITQDY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2148 TRHHPQQLSAPLPAPLYSFPGASCPVLDLRRPPSDLYLPPPDHGAPARGSPHSEGGKRSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2589 ERH----VSSMPAP-RQMKENQVPTKGTWRKIKESDISPTGMASQSASSGAASGAESKP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2643 L------IYQMAPPVSKTEDVWVRIEDCPI------NNPRSGRSPTGNT--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2268 PPAFFSKLTESNSAMVKSKKQEINKKLNTHNRNEPBYNISQPGTEIFNMPAITGTGLMTY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RSQAVQEHASTNMGLEAIIRKALMGKYDQWEESPPLSANAFNPLN----ASASLPAAMPI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-88328981; PubMed=1138108;
Lees J.F., Shneidman P.S., Skuntz S.F., Carden M.J., Lazzarini R.A.;
"The structure and organization of the human heavy neurofilament
subunit (NF-H) and the gene encoding it.";
EMBO J. 7:1947-1955(1988)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Homo sapiens (Human).
Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia, Euteria, Primates, Catarrhini, Hominidae, Homo.
NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NFH_HUMAN STANDARD, PRT; 1026 AA.
P12036, Q9UUS1;
01-0CT-1989 (Rel. 12, Catated)
10-0CT-2003 (Rel. 42, Last sequence update)
15-WAR-2004 (Rel. 43, Last annotation update)
Neurofilament triplet H protein (200 KDa neurofilament protein)
(Neurofilament heavy polypeptide) (NF-H).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Zhu Y., Han Y.;
"Molecular Cloning of human hSTE cDNA.";
Beijing Yi Ke Da Xue Xue Bao 31:531-531(1999)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RRIPLINRVWEDRPSSAGSTP 2461
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2820 SKTDITE-----SGAQSP 2833
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
TISSUE=Brain;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2090
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2765
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2441
                                                                                                                                              ò
                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY STANDARY S
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the BYML outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  31;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               --NIKINOAMRKKLILYFKRRNHARK---QWKOKFCORYDQ 304
                                                                                                                                                                                                               functioning
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |::|||| | | | | ::|
------RQREEARAARALARFAQEAEAARVDLQKKAQALQEECGYLRRHHQEEVGELL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         166 RLSKEELIQNMDRVDREITMVEQQISKLKKKQQQLEBERAAKPPEPEKPVSPPPIESKHRS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             226 LVQIIYDENRKKAEAAHRIL-----EGLGPQVEL-----PLYNQPSDTRQYHE----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       coil; Neurone; Phosphorylation; Repeat.
                                                                                                                                                                                                          PTM: Phosphorylation seems to play a major role in the functic of the larger neurofilament polypeptides (NF-M and NF-H), the levels of phosphorylation being altered developmentally and coincident with a change in the neurofilament function. SIMILARITY: Belongs to the intermediate filament family.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 225;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 1026;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                E -> A (IN REF. 1).
W; 0879B6A08D208C17 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COIL 2B.
MISSING (IN REF. 1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2.6%; Score 341.5; DB 1;
20.5%; Pred. No. 0.00017;
iive 147; Mismatches 405;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ROD.
TAIL.
COIL 1A.
LINKER 1.
COIL 1B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LINKER 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COIL 2A.
LINKER 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MIM; 162230; -.
GO; GO:0005883; C:neurofilament; NAS
GO; GO:0007399; P:neurogenesis; NAS
MEDLINE=99156230; PubMed=10048485;
                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; X15307; CAA33366.1; JOINED.
EMBL; X15308; CAA33366.1; JOINED.
EMBL; X15309; CAA33366.1; JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Intermediate filament; Coiled
                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; AF203032; AAF13722.1;
EMBL; AB020652; BAA74868.2;
PIR; S00979; QFHUH.
Genew; HGNC:7737; NEFH.
                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; X15306; CAA33366.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              112478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam; PF00038; filament; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                interPro; IPR001664; IF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PROSITE; PS00226; IF; 1
                                                                                                                                                                                               axonal caliber.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                811 87
1026 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Simi
Matches 200;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 269
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
g
```

..

us-09-522-753-5.rsp

```
InterPro; IPR008161; Clg_helix.
InterPro; IPR008160; Collagen.
InterPro; IPR000885; Fib_collagen (
InterPro; IPR002181; Fibrinogen_C.
InterPro; IPR001007; VWF_C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       X15896; CAA33904.1; -. M14423; AAA37333.1; -. M17491; AAA37334.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     U08020; AAA88912.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAA37332.1;
AAA37332.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAA37332.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     K03034; AAA37332.
K03035; AAA37332.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PIR; S57243; S21626.
MGD; MGI:88467; Collal.
                                                                                                                                                                                                                                 collagen protein.";
Gene 39:311-312(1985)
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              K03032;
K03033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          K03030;
                                                                                                                                  [2]
SEQUENCE OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL:
A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STANTANT AND A STAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             985 PREDA---APTKPAPPAPPPQNLQPESDAPQQPGSSPRGK-----SRSPA-PPA 1030
                                                          344 RQRSELEDRHQADIASYQEAIQQ-----LDAELRNTKW-----EMAAQLREYQ 386
                                                                                                                                                                                                                   -----QQSAELASMELNESSRWTE-----EEMETAKKGLLEHGRNWS 635
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 557 EAKSPPEAKSPEKEEAKSPAEVKSPEKAKSPAKEEAKSPAEAKSPEKAKSPVKEEAKSPA 616
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         636 AIARMYGSKTVSQCKNFYFNYKKRQNLDEILQQHKLKM------EKERNARRKKKAPA 688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EAKSPVKEEAKSPAEVKSPEKAKSPTKEEAKSPEKAKSPEKAKSPEKEEAKSPEKAKSPV 676
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                804
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   737 KTPEKAKSPVKEEAKSPEKAKSPEKAKTL--DVKSPEAKTPAKEBARSPADKFP--EKAK 792
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               193 SPVKEEVKSPEKAKSPL----KEDAKAPEKEIPKKEEVKSPVKEEEKPQEVKVKEPPKK 847
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EVDEAEGGDKWRLLSPRPSLLTPTGDPRANASPQKPLDLKQLKQLKQRAAAIPPIQVTKVHEP 984
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GOIOGSGAAQAQMQAETRDALKCDVTSALRE---IRAQLEGHAVQSTLQSEEWFRVRLDR 299
                                                                                                              AARSEHEVSEIIDGLSEQENLEKOMROLAVIPPMLYDADOORIKFINMNGLMADPMKVYK 424
                                                                                                                                                                                           DRQVMNMWSEQEKETFREKFMQHPK---NFGLIASFLERKTVAECVLYYYLTKKNENYKS 481
                                                                                                                                                                                                                                                                       LVRRSYRRGKSQQQQQQQQQQQQQQQQPMPRSSQEEKDEKEKEKEAEKEEKPEVEND 541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSSGAPQDSDSSATCSAD 924
                                                                                                                                                                                                                                                                                            :||: |: |: | : | : | : | : | EESTKSPPAEAASPEKEE-AKSPAEVKSPEKAKSPAKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            677 KAEAKSPEKAKSPVKAEAKSPVKEEAKSPVKEEAKSPVKEEA
                                      LMEALEKKVER I ENNPRRRAKESKVREYYEKOPPETRKORELOERMOSRVGORGSGLSMS
                                                                                                                                                                                                                                                                                                                                                  542 KEDLLKEKTDDTSGEDNDEKEAVASKGRKTANSQGRRKGRITRSMANEANSEEAITP---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ---AASBEAAFPPVVEDEEMEASGVSGNEEEMVEEAEALHASGNEVPRGECSGPATVNNS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SDTESIPSPHTEAAKDTGQ-NGPKPPATLGADGPPPGPPTPPRRTSRAPIEPTPASEATG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APTPPPAPPSPSAPPPVVPKEEKEEETAAAPPVEEGEEQKPPAAEELAVDTGKAEEPVKS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ----AEEEKAPATPKTEEKKDSKKE---EAPKKEAPKPKVEEKKEPAVEKPKESKV----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-JUL-1989 (Rel. 11, Created)
01-JUL-1989 (Rel. 11, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Collagen alpha 1(1) chain precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT; 1453 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1031 DKEAFAAEAQKLPGDPP 1047
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EKAEKSSSTDOKDSKPP 1014
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CA11 MOUSE
                                                                                                                                                                                                                                                                     482
                                                                                                                                                                                                                                                                                                                                                                                        498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             617
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       865 ]
243
                                                                                                                                                                                         425
                                                                                                                                                                                                                                                                                                                                                                                                                               599
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             866
                                      305
                                                                                                                365
                                                                                                                                                                                                                                 387
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      689
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           805
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            848
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  925
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      897
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                746
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                          요
                                                                                                                                                                                                                                                                                                                                                                                                              ⋧
                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ል 🖁 છે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      용
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                      ઠે
                                                                                                              8
                                                                                                                                                                                         8 6
                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                              ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ठ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          名
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ሯ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                            chain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE OF 1442-1453 FROM N.A.

SEQUENCE OF 1442-1453 FROM N.A.

MEDLINE=88124276; PubMed=3340560;

MEDLINE=88124276; PubMed=3340560;

"Two mRNAs of mouse pro alpha 1(I) collagen gene differ in the size of the 3'-untranslated region.";

Nucleic A-untranslated region.";

Nucleic Sec. 16:773-773 (1988).

-!- FUNCTION: Type I collagen is a member of group I collagen (fibrillar forming collagen).

-!- FUNCTION: Trimers of an ealpha 2(I) and two alpha 1(I) chains.
-!- TISSUE SPECIFICITY: Forms the fibrils of tendon, ligaments and bones. In bones the fibrils are mineralized with calcium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          hydroxyapatite. Prolines at the third position of the tripeptide repeating unit (G-X-Y) are hydroxylated in some or all of the chains. SIMILARITY: Contains 1 VWFC domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Monson J.M., Friedman J., McCarthy B.J.;
"DNA sequence analysis of a mouse pro alpha 1 (I) procollagen gene:
evidence for a mouse B1 element within the gene.";
Mol. Cell. Biol. 2:1362-1371(1982).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE OF 715-878 AND 1005-1058 FROM N.A.
MEDLINE-83157109; PubMed-6219867;
Monson J.M., McCarthy B.J.;
Midentification of a Balb/c mouse pro alpha 1(I) procollagen gene:
evidence for insertions or deletions in gene coding sequences.";
DNA 1:59-69(1981).
                                                                            for the mouse pro alpha 1(I)
                                                                                                                                                                                                                          SEQUENCE OF 518-1128 FROM N.A.
MEDLINE-86137403; PubMed=3841523;
French B.T., Lee W.-H., Waul G.G.;
"Nucleotide sequence of a cDNA clone for mouse pro alpha 1(1)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ပ
                                Li S.W., Khillan J., Prockop D.J.;
"The complete cDNA coding sequence
of type I procollagen.";
Matrix Biol. 14:599-595 (1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 [3]
SEQUENCE OF 735-1130 FROM N.A.
MEDLINE-83141374; PubMed-6298597;
MEDLINE=96033240; PubMed=8535610;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      X06753; CAA29927.1; -. K03036; AAA37332.1; -. CK03029; AAA37332.1; JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            K03031; AAA37332.1; JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAA37332.1; JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     JOINED
```

```
TRIIGEDSPSRL-DRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPH 1302
                                                                                                                                                                                                                                            1362 SYVEAQEDYLRREAKLLKREGTPPPPPRRDLTEAYKTQALGPL--KLKPAHEGLVATVK 1419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1473 --VRSLIGSPGR-----TFP--PVHPLDVMADARALERACYEESLKSRPGTASSSGGSI 1522
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1523 ARGAPVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSPVTWREPTPRLQEGSLSSSKASQD 1582
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1583 RKLTSTPREIAKSPHSTVPEHHPHPI----SPYEHLLRGVSGVDLYRSHIPLAFDPTSIP 1638
                                                                                                                                 ETAAPKRIYDMMEGRVGRAISSASIEGLMGRAIPP-ERHSPHHLKEQHHIRGSITQGIPR 1361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1639 RGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAALENROTIINDYITSQQMHHN 1698
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TATAMAQRADMLRGLSPRESSLALNYAAGPRGIIDLSQVPHLPVLVPPTPGTPATAMDRL 1758
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1848 SSRPASHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGIITAVEPSKPTVLRSTSTSSPV 1907
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----AGPQGPRGDKGETGEQGDRGIK 1097
                                                                                                                                                                                                                                                                                                    DLGAPGPSGARGERGFPGERGVOGPPGPA------GPRGNNGAPGNDGAKGDTG 702
                                                                                                                                                                                                                                                                                                                                                                                                                    AGPKGADGSPGKDGARGLTGPIGPPGPAGAPGD------KGEAGPSGPPGPTG 783
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ------GKGPRGET-----GPAG------RPGEVGP-PGPPGERG 924
                                                                            ---EAGAQGAPGPAGPAG
                                                                                                                                                                                       ------RGEQGPA-GSPGFQGLPGPAGPPGEAGKPGE------------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                : | | | | : | | : | | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | September | Septembe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GATGFPGAAGRVGP----PGPS-GNAGPP----GPPGPVGKEG--------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ------GPTHLTKPTTTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1793 SSERERDRDRERDRDREREKSILTSTTTVEHAPIWRPGTEQSSGSSGSSG-----GGGGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1016 DGAPGAKGDRGETGPAGPPGAPGAPGAPGPVGPAGKNGDRGETGPAGP-----AGPI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RPAATFPPATHCPLGGTLDGVYPTLMEPVLLPKEAPRVARPERPRADTGH------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1098 GHRGFSGLQGPPGSPGSPGEQGPSGASGPAGPRGPPGSAGSPGKDGLNGLPGPIGPPGPR
                                                                                                                                                                                                                                                                                                                                                            1420 BAGRSIHEIPREELRHTPELPLAPRPLKEGSITQGTP----LKYDTGASTTGSKKHD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Rattus norvegicus (Rat).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q9JIJJ4; Q9WUY7; Q9WW47; 28-FEB-2003 (Rel. 41, Created) 28-FEB-2003 (Rel. 41, Created) 28-FEB-2003 (Rel. 41, Last sequence update) 10-OCT-2003 (Rel. 42, Last annotation update) SH3 and multiple ankyrin repeat domains 3 (Shank3) (Proline-rich synapse associated protein 2) (ProSAP2) (SPANK-2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GRIGDSGPAGPPGP-PGPPGPPGPPSGGYDFSFLPQPPQEKSQ 1199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1991 HATIARTPAKNLAPHHASPDPPAPASASD-----PHREKTQ 2027
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ----AGPPGPAGPAGPPGPIGNVGAPGPKGPRGAAGP---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1759 AYLPTAPQPFSSRHSSSPLSPG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1815 AA
                                                                      ---AGE--PGKAGERGLPG--PPGAVGPAGKDG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1068 GPAGARGP------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STANDARD;
                                                                                                                                                                                          齒
                    1244
                                                                                                                                                                                       619
                                                                                                                                 1303
                                                                                                                                                                                                                                                                                                       655
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              737
                                                                         578
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  784
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         861
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      895
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 978
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               827
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SHK3 RAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SHANK3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 63
SHK3 RAT
ID SHK3
AC 09JI
DT 28-
DT 28-
DT 10
DE 5:
GN 0S
OC
                                                                                                                                                                                                                                                                                              a
                                                           셤
                                                                                                                         Š
                                                                                                                                                                              셤
                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PLPLGLHDTARPVLPRPPTISNPP-PLISSAKH------PSVL----ERQIGAISQGM 1129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      80;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EPPREDAAPTKPAPPAPPP------PQNLQPESDAPQ----QPGSSPRGKSRS 1025
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PA--PPADKEAFAAEAQKLPGDPPCWTSGLP-FPVPPREVIKASPHAPDPSAFSYAPPGH 1082
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AQEASVLRGI----ALGSVPGGSITKGIPSTRVPSDSAITYRGSI-THGTPADVLYKGTI 1243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               754
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         782
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PGPP------GPPGPPGLGGNFASQMSYGYDEKSAGVSVPGPMGPSGPR 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSSGAPQDSDSSATCSADEV 926
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DEAEGGDKNRLLSP----RPSLLTPTGDPRANASPQKPLDLKQLKQRAAAIPPIQVTKVH 982
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ----AGPKGEPGSPGENGAPGQMGPRGLPGERGRPGPP-----GTAGA-----RGN 313
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DGAVGAÁGPPGPTGPTGPPGRAGAKGEAGPÓGAR-GSEGPÓGVRGEPGPPGPAGAAG 372
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PAGNPGADGQPGAKGANGAPG----IAGAPGFPGARGPSGPPGPSGPPGPKGNSGEPGA 427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PGNKG--DTGAKGEPGATGVOGPPGPAGEEGKRGARGEPGPSGLPGPPGERG-GPGSRGF 484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----PGADGVAGPKGPSGERGAPGPAGPKG------SP-GEAGRPGEAGLPG 525
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      707 ASGVSGNEEEMVEEAEALHASGNEVPRGE------C-SGPATVNNSSDTESIPSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HTEA-----PPATLGADGPP--PGP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ------PAPPSPSAP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ----PPVVPKEEKEBETAAAPPVEEGE------EQKPPAAEELAVDTGKAEEPVKSEC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                177 GLPGPPGAPGPQGFQ----GPPGBPGGPGGSGPMGPRGPPGFNGDDGEAGKPGR---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SVQLHVPYSEHAKAPVGPV-TMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGVPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             526 AKGLIGSPGSPGPDGKIGPPGPAGDGRPGPAGPPGA----RGQAGVMGFPGP---KGT-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
CELL ATTACHMENT SITE (POTENTIAL).
CELL ATTACHMENT SITE (POTENTIAL).
A -> V (IN REF. 5).
Pfam; PF01410; COLFI; 1.
Pfam; PF01301; Collagen; 18.
ProDom; PD002007; Clg_Helix; 1.
ProDom; P0002078; Fib_collagen_C; 1.
SMART; SM00014; VMC; 1.
SMART; SM00214; VMC; 1.
PROSITE; PS01208; VWFC 1; 1.
PROSITE; PS01208; VWFC 1; 1.
PROSITE; PS01208; VWFC 2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 341; DB 1; Length 1453; Pred. No. 0.00025; Pred. No. 89; Mismatches 539; Indels 580;
                                                                                                                                                                                                                                                                                                                                                                                                                                         NONHELICAL REGION (N-TERMINAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                             TRIPLE-HELICAL REGION.
NONHELICAL REGION (C-TERMINAL)
                                                                                                                                                                                                                                                                                                                         AMINO-TERMINAL PROPEPTIDE.
COLLAGEN ALPHA 1(I) CHAIN.
CARBOXYL-TERMINAL PROPEPTIDE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A -> V (IN KEF. J., MW; 3B802E535DF81808 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                --PGÉRGPPGPQGARGLPGTA--GLPGMKGHRGFSGLDGAKGDAGP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PTPPRRTSRAPIEPTPASEATGAPTPP-----
                                                                                                                                                                                                                                                                   Collagen; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 21.7%;
Matches 335; Conservative 8:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AA; 137944
                                                                                                                                                                                                                                                                                                                                                1207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                11181
1207
56
1354
736
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1450
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1450
                                                                                                                                                                                                                                                                Glycoprotein;
SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   755
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                783
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           819
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        867
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        927
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        983
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1026
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1083
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     428
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  485
                                                                                                                                                                                                                                                                                                                                             CHAIN
PROPEP
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
```

ò 셤 à පු ò 임 ઠ g ò g ò d

g ò

ò Ωp ð g à 셤 ð

1792

977

1847

1015

1957

```
to 1000 due to a frameshift.
                                                                                                                                                                                                                                                                 InterPro; IPR001478; PDZ.
InterPro; IPR001660; SAM.
InterPro; IPR001452; SH3.
                                                                                                                                                                                                                                                                                                                                                                                                    PD000066; SH3; 1.
                                                                                                                                                                                                                                                 IPR002110; ANK
                                                                                                                                                                                                                                                                                                                       Pfam; PF00023; ank; 5.
Pfam; PF00595; PDZ; 1.
Pfam; PF00536; SAM; 1.
Pfam; PF00018; SH3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1492
1590
1815
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1815
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1386
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         11115
11162
11167
11337
11348
11348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         781
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ANK repeat; SH3 domai
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                914
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1612
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1386
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1345
1348
1354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1570
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1212
                                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                      ProDom;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MUTAGEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MUTAGEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
 Sheng M., Kim E.;
"The Shank family of scaffold proteins.";
"The Shank family of scaffold proteins.";
J. Cell Sci. 113:1851-1856(2000)
J. Cell Sci. 113:1851-1856(2000)
J. Cell Sci. 113:1851-1856(2000)
J. Cell Sci. 113:1851-1856(2000)
J. Cell Sci. 113:1851-1856(2000)
J. Cell Sci. 113:1851-1856(2000)
J. Cell Sci. 113:1851-1856(2000)
J. Cell Sci. 113:1851-1856(2000)
J. Cell Sci. 113:1851-1856(2000)
J. Cell Sci. 113:1851-1856(2000)
J. Cell Sci. 113:1851-1856(200)
J. Cell Sci. 113:1851-1856(200)
J. Cell Sci. 113:1851-1856(200)
J. Cell Sci. 113:1851-1856(200)
J. Cell Sci. 113:1851-1856(200)
J. Cell Sci. 113:1851-1856(200)
J. Cell Sci. 113:1851-1856(200)
J. Cell Sci. 113:1851-1856(200)
J. Cell Sci. 113:1851-1856(200)
J. Cell Sci. 113:1851-1856(200)
J. Cell Sci. 113:1851-1856(200)
J. Cell Sci. 113:1851-1856(200)
J. Cell Sci. 113:1851-1856(200)
J. Cell Sci. 113:1851-1856(200)
J. Cell Sci. 113:1851-1856(200)
J. Cell Sci. 113:1851-1856(200)
J. Cell Sci. 113:1851-1856(200)
J. Cell Sci. 113:1851-1856(200)
J. Cell Sci. 113:1851-1856(200)
J. Cell Sci. 113:1851-1856(200)
J. Cell Sci. 113:1851-1856(200)
J. Cell Sci. 113:1851-1856(200)
J. Cell Sci. 113:1851-1856(200)
J. Cell Sci. 113:1851-1856(200)
J. Cell Sci. 113:1851-1856(200)
J. Cell Sci. 113:1851-1856(200)
J. Cell Sci. 113:1851-1856(200)
J. Cell Sci. 113:1851-1856(200)
J. Cell Sci. 113:1851-1856(200)
J. Cell Sci. 113:1851-1856(200)
J. Cell Sci. 113:1851-1856(200)
J. Cell Sci. 113:1851-1856(200)
J. Cell Sci. 113:1851-1856(200)
J. Cell Sci. 113:1851-1856(200)
J. Cell Sci. 113:1851-1856(200)
J. Cell Sci. 113:1856(200)
J. Cell 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TISSUE SPECIFICITY: Predominantly expressed in brain.
SIMILARITY: BELONGS TO THE SHAWK FAMILY.
SIMILARITY: Contains 4 ANK repeats.
SIMILARITY: Contains 1 Brerile alpha motif (SAM) domain.
SIMILARITY: Contains 1 Starile alpha motif (SAM) domain.
SIMILARITY: Contains 1 SH3 domain.
SIMILARITY: Contains 3 SH3 domain.
SIMILARITY: Contains 1 SH3 domain.
SIMILARITY: Contains 1 SH3 domain.
SIMILARITY: Contains 1 SH3 domain.
CAUTION: Ref.2 sequence differs from that shown in postitions 3 to 37 due to a frameshift.
CAUTION: Ref.3 sequence differs from that shown in postitions 973
                                                                   MEDLINE-99458653; PubMed=10527873;
Boeckers T.M., Winter C., Smalla K.-H., Kreutz M.R., Bockmann J.,
Seidenbecher C., Garner C.C., Gundelfinger B.D.;
Proline rich synapse associated proteins ProSAP1 and ProSAP2 interact
with synaptic proteins of the SAPAP/GKAP family.";
Biochem. Biophys. Res. Commun. 264:2476-2528(1999).
                                                                                                                                                                                                                                                                                                           t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MUTAGENESIS LEU-1386 AND CYS-1389.
MEDLINE=99360651; PubMed=1043269;
TU J.C., Xiao B., Naisbitt S., Yuan J.P., Petralia R.S., Brakeman P., Doan A., Aakalu V.K., Lanahan A.A., Sheng M., Worley P.F.;
"Coupling of mGluk/Homer and PSD-95 complexes by the Shank family of popersynaptic density proteins.";
Neuron 23:583-592(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          and synaptic junction.
SUBUNIT: May homomultimerize via its SAM domain (By similarity).
Interacts with DLGAPI/GKAP, MGLURIA and MGLURS C-termini via its PDZ domain. Interacts with Homer-1, Homer-2, Homer-3 and CCTN/cortactin SH3 domain. Is part of a complex with DLG4/PSD-95 and DLGAPI/GKAP.
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
                                                                                                                                                                                                                                                                                                                                                                                                                                    Tobaben S., Suedhof T.C., Stahl B.; "The G protein-coupled receptor CL1 interacts directly with proteins of the shank family.";
                                                                                                                                                                                                            SEQUENCE FROM N.A. (ISOFORM 2), AND INTERACTION WITH DLGAP1 AND CCTN STRAIN=Sprague-Dawley; TISSUE=Hippocampus;
MEDLINE=99360650; PubMed=10433268;
                                                       SEQUENCE FROM N.A. (ISOFORM 1), AND INTERACTION WITH DIGAP1 AND DIG4
                                                                                                                                                                                                                                                             Naisbitt S., Kim B., Tu J.C., Xiao B., Sala C., Valtschanoff J., Weinberg R.J., Worley P.F., Sheng M.; Shank, a novel family of postsynaptic density proteins that binds the NMDA receptor/PSD-95/GKAP complex and cortactin."; Neuron 23:569-582(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SUBCELLULAR LOCATION: Cytoplasmic; postsynaptic density of neuronal calls. Extends into the region subjacent to the PSD. ALTERNATIVE PRODUCTS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         INTERACTION WITH HOMER PROTEINS; DLGAP1; MGLURIA AND MGLURS, AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IsoId=09JLU4-3; Sequence=VSP_006087, VSP_006088;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Event=Alternative splicing; Named isoforms=3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IsoId=Q9JLU4-2; Sequence=VSP_006089;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        IsoId=Q9JLU4-1; Sequence=Displayed;
                                                                                                                                                                                                                                                                                                                                                                                  PARTIAL SEQUENCE FROM N.A. (ISOFORM 3)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Biol. Chem. 275:36204-36210(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE=20549637; PubMed=10958799;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE=20267867; PubMed=10806096;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Name=2; Synonyms=A;
                                                                                                                                                                                                                                                                                                                                                                                                        TISSUE=Brain;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Name=1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Name=3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REVIEW
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the BMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                P->L: ABOLISHES INTERACTION WITH HOMER-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      F->C: ABOLISHES INTERACTION WITH HOMER-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -> PRRRAGMV (in isoform 3).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IN REF. 2).
IN REF. 2).
IN REF. 3).
IN REF. 3).
IN REF. 3).
IN REF. 3).

SHIGHQ (IN REF. 3).
                                                                                                                                                                                                                                                                                                                                                                                            SWART; SW00248; ANK; 5.
SWART; SW00228; PDZ; 1.
SWART; SW00454; SAM; 1.
SWART; SW00326; SH3; 1.
PROSITE; PS50088; ANK REPEAT; 4.
PROSITE; PS50005; ANK REP_REGION; 1.
PROSITE; PS50005; SAM DDZ; 1.
PROSITE; PS50002; SAM DOMAIN; 1.
PROSITE; PS50002; SH3; 1.
ANK repeat; SH3; domain; SH3-binding; Coiled coil; Repeat;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ANK 1.
ANK 2.
ANK 3.
ANK 4.
COILED COIL (POTENTIAL).
SH3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (In isoform 1). P 006089.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /FTId=VSP 006087.
Missing (In isoform 3).
/FTId=VSP 006088.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (IN REF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SH3-BINDING
COILED COIL
                                                                                                                               or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ZI)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ^ ^
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FTIG=VSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SPTPVHSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MQ -- AA
                                                                                                                                                                                     EMBL; AF133301; AAF61375.1; ALT_FRAME
EMBL; AF159047; AAD42976.1; ALT_FRAME
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Missing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Missing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ISOFORM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ISOFORM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ტ
^ ^
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                P -> Y
                                                                                                                                                                     EMBL; AJ133120; CAB45688.1;
```

us-09-522-753-5.rsp

۶ ۶	1550 FAGHLPRGSPV	н 160
a ;	201	TT
<u>-</u> ہ	604 HPHPISPY	RGVSGVDLYRSHIPLAFDPTSIPRGIPLDAAAAYYLPR
Д		ERLLGTGATTGRDLLLPSPVSALKPLVGG 1160
×	1662 PHLYPPYLIRG	QRADMLR(
Д	1161 ÞSLGÞ	SGSTFIHP-LTGKPLDPSSPLALALAARERALA 1197
٨	1722 LNYAAGPRGIII	LNYAAGPRGIIDLSQVPHLPVLVPPTPGTPATAMDRLAYLPTAPQPFSSRHS 1773
Д	1198	SQTPSRSPTPVHSPDADRPGPLFVDVQTRDSER 1230
>-	1774SSPLSPG	TKPTTTSS
Д	1231 GPLASPAFSPRSPAWIPVP	
>-	1829 PGTEQSSGSSG	PGTEQSSGSSGSSGGGGGSSSRPASHSHAHQHS-PISPRTQDALQQRPSVLHNTGMKGII 1887
Ω	1279	SLQRPAGLIVVHATSNGQEPNRLGAEEERP 1308
>-	1888 TAVEPSKPTVL	TAVEPSKPTVLRSTSTSSPV-RPAATFPPATHCPLGGTLDGVYPTLMEPVLLP 1939
۵	1309 -GTPELAPTPM	GTPELAPTPWQAAAVAEPMPSPRAQPPGSIPADPGPGGSSE-EEPELVFAVNLPPAQL 1366
>-	1940KBAI	KEAPRVARPERPRADIGHAFLAKPPARSGLEPASSPSKGS-EPRPLVPPVSG 1990
۵	: 1367 SSSDEETREELA	SDEETREELARIGLVPPPEEFANGILLATPPPGPGPLPTTVPSPASGKPSSELPP 1423
>	1991 HATIARTPAKNLAPHHASPD	APHHASPDPPAPPASASDPHREKTQSKPFSIQELELRSLGYHGSSY 2048
٥	1424	: APESAADSGVEEADTRSSSDPHLETTST1451
_	2049 SPEGVEPVSPVSSPS	-LTHDKGLPK
0	1452ISTVS	: : : : : TLSSESGELTDTHTSFADGHT
_	2106 LR-PLPESQPS-	VKGHQRVVTLAQHISEVITQD
0	1494 LKSPLGKGPVTE	FRDPLLKQSSDSELMAQQHHATSTGLTSAAGP 1536
_	2161 A-PLYSFPGASC	A-PLYSFPGASCPVLDLRRPPSDLYLPPPDHGAPARGSPHSEGGKRSPEPNKTSVLGGGE 2219
0	1537 ARPRYLFQ	1
_	2220 DGIEPVSPPEGN	DGIEPVSPPEGMTEPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQPPAFFSKLTESN 2279
0	1571	1570
_	2280 SAMVKSKKQEIN	KLNTHNRNEPEYNISOPGTEIFNMPA
0	1571 -SELSSRLQQLNK	KDTRSLGEBPVGGLGSL1598
	2340 MGLEAIIRKALM	FNPLNA
•	1599	LDPAKKSPIAAARCAVVPSAGWLFSSLGELS-TISAQRS 1636
	2397 PGG-GGKAKVSGRPS	GLASGDRPPS
•	1637 PGGPGGGASYSV	PGGPGGGASYSVRPSGRYPVARRAPSPVKPASLERVEGLGAGVGGAGRPFGLT 1689
	2447 NRVWEDRPSS	-SAGSTPFPYNPLIMRLQAGVMASPPPFGLPAGSGPLAGPHHA 2497
	1690PPTILK	PPTILKSSSLSIPHEPKEVRFVVRSVSARSRSPSPSPLPSPSPGSGPSAGPRRP 1743
	2498 WDEEPKPL 250	S.
•	1744 FQQKPLQL 1751	1

```
THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SECTION OF THE SE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 "Structure of cDNA clones coding for the entire prepro alpha 1 (III) chain of human type III procollagen. Differences in protein structure from type I procollagen and conservation of codon preferences."; Blochem. J. 260:509-516(1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                "Covalent structure of collagen: amino acid sequence of cyanogen
bromide peptides from the amino-terminal segment of type III collagen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE OF 399-727.

MEDLINE=79000343; PubMed=687591;
Seyer J.M., Kang A.H.;
"Covalent structure of collagen: amino acid sequence of five consecutive CNBr peptides from type III collagen of human liver.";
Biochemistry 17:3404-3411(1978).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE OF 950-1466 FROM N.A.
MEDIINE=88189827; PubWed=3357782;
Mankoo B.S., Dalgleish R.;
MHANDO B.S., Dalpha 1 (III) collagen: cDNA sequence for the 3' end.";
Nucleic Acids Res. 16:2337-2337(1988).
                                                                                                                                                                                                                                                                                      Craniata, Vertebrata, Buteleostomi,
Catarrhini, Hominidae, Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           the entire human alpha 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            artefact.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDICANES 01,75,701.
MEDICANES 198282; PubMed=6246925;
Seyer J.M., Mainardi C., Kang A.H.;
"Covalent structure of collagen: amino acid sequence of alpha 1 (III)-CBS from type III collagen of human liver.";
Biochemistry 19:1583-1589(1980).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE OF 1065-1466 FROM N.A.
MEDLINE=85087944; PubMed=6096827;
Loidl H.R., Brinker J.M., May M., Pihlajaniemi T., Morrow S.
Rosenbloom J., Myers J.C.;
                                                                                                                                                                                                                                                                                                                                                                                                                                    TISSUE=Skin fibroblast;
MEDLINE=89350838; PubMed=2764886;
Ala-Kokko L., Kontusaari S., Baldwin C.T., Kuivaniemi H.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a cDNA cloning
                                                                        POLATE INTERPRETATION TO THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE P
                                                         1466 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Submitted (DEC-1977) to the PIR data bank
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Janeczko R.A., Ramirez F.; "Nucleotide and amino acid sequences of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Molyneux K., Dalgleish R.; "Human type III collagen 'variant' is a Nucleic Acids Res. 16:11833-11833(1988)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (III) collagen.";
Nucleic Acids Res. 17:6742-6742(1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE=89386015; PubMed=2780304;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE=89098346; PubMed=3211760;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE=77134724; PubMed=557335;
Seyer J.M., Kang A.H.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Siochemistry 16:1158-1164(1977)
                                                                                                                                                                                                                                                                                      Eukaryota; Metazoa; Chordata;
Mammalia; Eutheria; Primates;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE OF 149-1225 FROM N.A
                                                      STANDARD;
                                                                                                                                                                                                                                                              Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE OF 168-398.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE OF 728-964.
                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REVISION TO 1184.
                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         human liver
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Prockop D.J.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Seyer J.M.;
                                                      HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REVISIONS
                                                                                                                                                                                                                                   COL3A1.
                           CA13_HUMAN
```

MEDLINE-97255959; PubMed-9101290; Kuivaniemi H., Tromp G., Prockop D.J.; Wutations in fibrillar collagens (types I, II, III, and XI), fibrillassociated collagen (type IX), and network-forming collagen (type X) cause a spectrum of diseases of bone, cartilage, and blood vessels."; MEDLINE=93299988; PubMed=8514866; Tromp G., Mu Y., Prockop D.J., Madhatheri S.L., Kleinert C., Earley J.J., Zhuang J., Noerrgaard O., Darling R.C., Abbott W.M., Cole C.W., Jaakkola P., Ryynaenen M., Pearce W.H., Yao J.S.T., Majama K., Smullens S.N., Gatalica Z., Ferrell R.E., Jimenez S.A., Jackson C.E., Michels V.V., Kaye M., Khivaniemi H.E., Jimenez S.A., "Sequencing of cDNA from 50 unrelated patients reveals that mutations in the triple-helical domain of type III procollagen are an infrequent cause of aortic aneurysms."; Clin. Invest. 91:2539-2545(1993). VARIANT AORTIC ANEURYSM ARG-786. MEDLINE=91056145; PubMed=2243125; Kontusaari S., Tromp G., Kuivaniemi H., Romanic A.M., Prockop D.J.; "A mutation in the gene for type III procollagen (COL3A1) in a family 'Molecular cloning and carboxyl-propeptide analysis of human type III SEQUENCE OF 1176-1466 FROM N.A.
MEDLINE-85157600; PubMed=2579949;
Chu M.-L., Weil D., de Wet W.J., Bernard M.P., Sippola M., Ramirez F.
"Isolation of cDNA and genomic clones encoding human pro-alpha 1
(III) collagen. Partial characterization of the 3' end region of the Toman D., Ricca G., de Crombrugghe B.;
"Nucleotide sequence of a cDNA coding for the amino-terminal region
of human prepro alpha 1(III) collagen.";
Nucleic Acids Res. 16:7201-7201(1988). Zafarullah K., Kleinert C., Tromp G., Kuivaniemi H., Kontusaari S., Wu Y., Ganguly A., Prockop D.J.; "G to A polymorphism in exon 31 of the COL3A1 gene."; "Nucleic Acids Res. 18:6180-6180 (1990). SEQUENCE OF 1161-1200 FROM N.A.
MEDLINE=86187804; PubMed=3754462;
Miskulin M., Dalgleish R., Kluve-Beckerman B., Rennard S.I.,
Tolstoshev P., Brantly M., Crystal R.G.;
"Human type III collagen gene expression is coordinately modulated
with the type I collagen genes during fibroblast growth.";
Biochemistry 25:1408-1413(1986). Seyer J.M., Kang A.H.; "Covalent structure of collagen: amino acid sequence of alpha SEQUENCE OF 1-176 FROM N.A.
MEDLINE-89378752; PubMed-2777083;
Benson-Chanda V., Su M.W., Weil D., Chu M.-L., Ramirez F.;
Benson-Chanda V., Su M.W., Weil D., Chu M.-L., Ramirez F.;
"Cloning and analysis of the 5' portion of the human type-III procollagen gene (COL3A1)."; VARIANT AORTIC ANEURYSM ARG-303, AND VARIANT THR-668. 1(III)-CB9 from type III collagen of human liver."; Biochemistry 20:2621-2627(1981). Nucleic Acids Res. 12:9383-9394 (1984) Biol. Chem. 260:4357-4363(1985) VARIANT THR-698. MEDLINE=91045136; PubMed=2235526; MEDLINE=88303360; PubMed=3405773; MEDLINE=81208139; PubMed=7016180; Hum. Mutat. 9:300-315(1997). SEQUENCE OF 1-170 FROM N.A. SEQUENCE OF 965-1200. 78:255-265 (1989) REVIEW ON VARIANTS

e

```
ઠે
                                                                                      g
                                                                                                                      ò
                                                                                                                                                                          ò
                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                              셤
                                                                                                                                             셤
                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 82;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DDIICDDQELDCPNPEIPFGECCAVC-----PQPPTAPTRPPNGQGPQGPKGDPGPP 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                771 ATLGADGPP------PGPPTPPRRTSRAPIEP---TPASEA------TGAPT 807
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   P-----PPAPP-----SPSAPPVVPKEEKEBETAAAPPVEEGEEQKPPAAEELAVDT 855
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          715 EEMVEEAEALHASGNEVPRGECSGPATVNNSSDTESIPSPHTEAAKDTGQNGPK----PP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     175 PAGPPGPPGPPGTSGHPGSPGSPGYQGPPGEPGQAGPSGPP-----GPPGA-----
                                                                                                                                                                                                                                                                                                                                                                                                                       Pope F.M.;
"A single base mutation in the gene for type III collagen (COL3A1)
converts glycine 847 to glutamic acid in a family with Enlers-Danlos
syndrome type IV. An unaffected family member is mosaic for the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VARIANT EDS-IV ASP-1050.
MEDLINE=90037070; PubMed=2808425;
Tromp G., Kuivaniemi H., Stolle C.A., Pope F.M., Prockop D.J.;
"Single base mutation in the type III procollagen gene that converts the codon for glycine 883 to aspartate in a mild variant of Ehlers-Danlos syndrome IV.";
J. Biol. Chem. 264:19313-19317(1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                 Tromp G., Kuivaniemi H., Shikata H., Prockop D.J.;
"A single base mutation that substitutes serine for glycine 790 of the alpha 1 (III) chain of type III procollagen exposes an arginine and causes Ehlers-Danlos syndrome IV.";
J. Biol. Chem. 264:1349-1352(1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pope F.M.;
"Characterisation of a glycine to valine substitution at amino acid position 910 of the triple helical region of type III collagen in a patient with Ehlers-Danlos syndrome type IV.";
J. Med. Genet. 28:458-463(1991).
                  VERTANT EDS-IV ARG-828.
WICHDLINE-94016385; PubMed=8411057;
Richards A.J., Narcisi P., Lloyd J.C., Ferguson C., Pope F.M.;
Withe substitution of glycine 661 by arginine in type III collagen produces mutant molecules with different thermal stabilities and J. Med. Genet. 30:690-693(1993).
                                                                                                                                                                                                                                                             VARIANT EDS-IV VAL-960.

MEDLINE-95268429; PubMed=7749417;
Tromp G., de Raepe A., Nuytinck L., Madhatheri S.L., Kuivaniemi "Substitution of valine for glycine 793 in type III procollagen Ehlers-Danlos syndrome type IV.";
Hum. Mutat. 5:179-181(1995).
                                                                                                                                                                                                                                                                                                                                                                           VARIANT EDS-IV GLU-1014.
MEDLINE-92316511; Pubmed-1352273;
Richards A.J., Ward P.N., Narcisi P., Nicholls A.C., Lloyd J.C.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2.6%; Score 340; DB 1; Length 1466;
21.1%; Pred. No. 0.00027;
tive 89; Mismatches 600; Indels 648;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE=91374480; PubMed=1895316;
Richards A.J., Lloyd J.C., Ward P.N., de Paepe A., Narcisi P.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VARIANT EDS-IV GLU-1173.
MEDLINE-93022543; PubMed=1357232;
Johnson P.H., Richards A.J., Pope F.M., Hopkinson D.A.;
with aortic aneurysms.";
J. Clin. Invest. 86:1465-1473(1990)
                                                                                                                                                                          MEDLINE=89109135; PubMed=2492273;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Genet. 89:414-418(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches 357; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VARIANT EDS-IV VAL-1077
                                                                                                                                                           VARIANT EDS-IV SER-957
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         808
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Local
셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
```

<u>≽</u> •	856	GKAEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSSGAPQDS 915
≥ •	916	PSLLTPTGDPR-ANASPQKPLDLKQLKQRAAAIP
۵	237	
ء <	975	PIQVTKVHEPPREDAAPTKPAPPAPPPQNLQPESDAPQQPGSSPRGKSRSPA 1027
· >		OKEAFAAEAQKILPGDPPCWTSGLPFPVPPREVIKASPHAPDPSAFSYAPPGHPLPLG 10
Д	334	
>-	1088	LHDTARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEHAKAPVGP 1147
۵	376 1	HAGAQGPPGPPGINGSPGGKGEMGPAGIPGAPGLMGARGPPGP 418
> Q	1148	V-TWGLPLPWDPKKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEASVLRGT 1198
>-	1199 /	PSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPSRLDRG 12
Ω	474 (
>	1259	REDSL-PKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETAAPKRTYDM 1313
<u>م</u>	511 /	APGPAGPRGAAGEPGRDGVPGGPGMRGMPGSPGGPGSDGKPGPPGS 556
>-	1314	MEGRVGRAISSASIEGLMGRAIPPERH-SPHHLKEQHHIRGSITQGIPRSYVEAQ 1367
۵	557	-desgepegpegpegpgymgfpgpkgndgapgkngerggpggpgpgpgpgrngerg 615
> .		DYLRREAKLLKREGTPPPPPPSRDLTEAYKTQALGPLKLKPAHBGLVATVKEAGRSIHE :
Ω		PPGPPGPTGPGGDKGDTGPPGPQGLQGLPGTGGP 649
>-		TPELPLAPRPLKEGSITQGTPLKYDTGASTTGSKKHDVRSLIGSPGRTFP
۵	650	-PGENGKPGEPGPKGDAGAPCAPGKGDAGAPGERGPP- 686
> 4		SSSGGSIARGAPVIVPELGKPRQSPLTYEDHG 15
۵		GPEGGKGAAGPPG 714
~ 0	1548 A	APFAGHLPRGSPVTWREPTPRLQEGSLSSSKASQDRKLJTSTPREIAKSPHSTVPEHHPHP 1607
> :	1608 I	ISPYEHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPP 1667
	751 -	
	1668 Y	YLIRGYPDTAALENRQTIINDYITSQQMHHNTATAWAQRADMLRGLSPRESSLALNYAAG 1727
0	- 171	GPAGQPGDKGEGGAPGLPGIAG 792
~ (IIDLSQVPHLPVLVPPTPGTPA-TAMDRLAYLPTAPQPPSSRHSSSPLSPG 17
0		PRGSPGERGEIGPPGPAGFPGAPGNGEPGGKGERGAPGEKGEG 836
~ 0	1781 G 837 G	GPTHLTKPTTTSSSERERDRDRERDRDREREKSILTSTTTVEHAPIWRPGTEGSSGSS 1840
_	1841 S	SGGGGGSSSRPASHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGIITAVEPSKP 1895
•	866 P	
	1896 T	TVLRSTSTSSPVRPAATFPPATHCPLGGTLDGVYPTLMEPVLLPKEAPRV 1945

```
MUTAGEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Genew;
MIM; 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                           +
    DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŝ
    2209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2242
                                                                                                                                                                                                                                                                                                                                 2164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                1186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1187 PGPPGAPGPCCGGVGAAAIAGIGGEKAGGFAPYYGDEPMDFKINTDEIMTSLKSVNGOIE 1246
                                                               ARPERPRADTGHAFLAKPPARSGLEPASSPS----KGSEPRPLVPPVSGHATIARTPAKN 2001
                                                                                                        -----TGARGLAGPGMPG--PRGSPGPQGVKGESGKP------GANG
                                                                                                                                                                                 2002 LAPHHASPDP---PAPPASASDPHREKTQSKPFSIQELELRSLGYHGSSYSPEGV---E
                                                                                                                                                                                                                                          2055 PVSPVSSPSLTHDKGLPKHLEELDKSHLEGELRPKQPGPVKLGGEAAHLPHLRPLPESQP
                                                                                                                                                                                                                                                                                   1037 NGSP-GAPGAPGHPGPPGPVGPAGKSGDRGESGP--AGPAGAPGPAGS----RGAPGPQG
                                                                                                                                                                                                                                                                                                                                 2115 SSSPLLQT----APGVKGHQRVVTLAQHISEVITQDYTRHHPQQLSAP-LPAP----LY
                                                                                                                                                                                                                                                                                                                                                                            ------RGFPGNPGAPGSPGPAGQGGAI
                                                                                                                                                                                                                                                                                                                                                                                                                       2165 SFPGASCPVLDLRRP------PSDLYLPPPDHGAPARGSPHSEGGKRSPEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                1131 GSPGPAGP----RGPVGPSGPPGKDGTSGHPGPIGPPGPRGNRGERGSEGSPGHPGQPGP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -----VSPPEGMTEPGHSRAVY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3), PARTIAL SEQUENCE, FUNCTION, AND MUTAGENESIS OF ASP-1093.
MEDLINE-99418558; PubMed-10490026;
Sahara S., Aoto M., Eguchi Y., Imamoto N., Yoneda Y., Tsujimoto Y.;
"Acinus is a caspase-3-activated protein required for apoptotic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ishikawa K.-I., Nagase T., Suyama M., Miyajima N., Tanaka A.,
Kotani H., Nomura N., Ohara O.;
"Prediction of the coding sequences of unidentified human genes. X.
The complete sequences of 100 new cDNA clones from brain which can
code for large proteins in vitro.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q9UKV3: 075158; Q9UG91; Q9UKV1; Q9UKV2; 28-FEB-2003 (Rel. 41, Created)
28-FEB-2003 (Rel. 41, Last sequence update)
28-FEB-2003 (Rel. 41, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Apoptotic chromatin condensation inducer in the nucleus (Acinus).
ACINUS OR KIAA0670.
Homo sapiens (Human).
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TISSUE=Uterus;
Wambutt R., Heubner D., Mewes H.-W., Gassenhuber J., Wiemann S.;
Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
Li W.B., Gruber C., Jessee J., Polayes D.; "Pull-length cDNA libraries and normalization."; Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1341 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE OF 56-1341 FROM N.A. (ISOFORM 1)
                                                                                                                                                                                                                                                                                                                                                            1090 PRGDKGETGERGAAGIKGH-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE=98403880; PubMed=9734811;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A. (ISOFORM 4).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A. (ISOFORM 3).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |: ||: ||: |:|
1247 SLISPDGSRKNPAR 1260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PLLYRDGEQTEPSR 2256
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Nature 401:168-173(1999)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                condensation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                brain;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TISSUE=Fetal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACIN HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   chromatin
                                                                  1946
                                                                                                                                                                                               886
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2243
                         916
                                                                                                           953
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2210
                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                              ઠે
                                                                                                                                                                                                                                                                                                                                                                     Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                               ઠે
                                                                                                      셤
                                                                                                                                                 δ
                                                                                                                                                                                        g
                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HERE REPORTED THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SERVICE OF THE SER
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 63;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GO; GO:0019899; F:enzyme binding; NAS.
GO; GO:0003676; F:nucleic acid binding; NAS.
GO; GO:00030263; P:apoptotic chromosome condensation; IDA.
GO; GO:0030218; P:erythrocyte differentiation; IBP.
GO; GO:0045657; P:positive regulation of monocyte differentia. . . .; IBP.
InterPro; IPR003034; SAP.
                                                                                                                                                                                                                                                                        Isold=Q9UKV3-4; Sequence=VSP 004027;
TISSUE SPECIFICITY: Ubiquitous—
PTW: Undergoes proteolytic cleavage; the processed form is active,
contrary to the uncleaved form.
SIMILARITY: Contains 1 SAP domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -> MSPADRCKSANTIEPATTSSLÄLFLLLQRDQSSRTRGL
P (in isoform 2)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /FTId=VSP_004027.
GSPKKCEAEEAEPPAATQPQTSETQTSHLPESERIHHTV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /FTIG=VSP 004029.
D->A: ABOLISHES CLEAVAGE BY CASP3 AND CHROMATIN CONDENSATION ACTIVITY.
Q -> H (IN REF. 4)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /FTId=VSP_004028.
SERIHHTV -> MLSESKEG (in isoform 3).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2.6%; Score 339.5; DB 1; Length 1341; ilarity 18.8%; Pred. No. 0.00026; Conservative 210; Mismatches 608; Indels 557;
                    FUNCTION: Induces apoptotic chromatin condensation after
                                                                                                                                                                                           Sequence=VSP_004025, VSP_004028;
                                                                                                                                                                                                                                      IsoId=Q9UKV3-3; Sequence=VSP_004026, VSP_004029;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ARG/ASP/GLU/LYS-RICH.
CLEAVAGE (BY CASPASE-3)
Missing (in isoform 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Missing (In isoform 4).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Missing (in isoform 3).
/FIId=VSP 004026.
                                    Alternative splicing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     004025
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GO:0005634; C:nucleus; IDA.
GO:0016887; F:ATPase activity; NAS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /FTIG=VSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SER-RICH.
PRO-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GLU-RICH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; AF124726; AAD56724.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL, AF124728; AAD56726.1; -- EMBL, ALOS0382; CAB43681.1; -- EMBL; BX247975; CAD62309.1; -- EMBL; AB014570; BAA31645.2; --
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           151887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HGNC:17066; ACINUS
Res. 5:169-176(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PROSITE; PS50800; SAP; 1
                                                                                                                                                                   Name=2; Synonyms=S';
IsoId=Q9UKV3-2; Se
                                                                                                                                                                                                                Name=3; Synonyms=S;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1093
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1131
1341
1094
727
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           758
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               766
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   766
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam; PF02037; SAP; 1.
SMART; SM00513: SAP: 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         139 13
1341 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SM00513; SAP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Apoptosis; Nuclear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity
Matches 319; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1114
1132
1093
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1093
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               728
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   759
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /ARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CONFLICT
```

1286 VTQCSKEDGRSSSGPPHE	1459 1050 1510 1101 1567 1161	DD 1213DLFRK	RESULT 66 CA44 HUMAN ID	NCBI_TaxID=9666; [1] SEQUENCE FROM N.A. TISSUE=Kidney; MEDLINE=55014445; Pr Leinonen A., Marriyat "Complete primary si "Complete primary si "Complete primary si "Complete primary si "Complete primary si "Complete primary si "Complete primary si "Complete primary si "Complete primary si "Sequence of primary si [2] SEQUENCE OF 1-23 FRC MEDLINE=9819684; Pr MOMOCTA R., SUGIMOCO "The primary si "The
OYHENIKINOAMR-KKLILYFKRRNHARKQWKQKFCQRY	KKNENYKSLVRRSYRRGKSQQQQQQQQQQQQQQDPRSSQ : : :	374		1061 EVIKASPHAPDFSAFSTAPPGHPLPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLER 1120 655SLSPGVSRDSSTSYTETKDPSGGEVATPPVPQLQVCEPKERTSTSSSSVQAR 707 1121 QIGAISQGMSVQLHVPYSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRG 1176 1121 QIGAISQGMSVQLHVPYSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRG 1176 1121 QIGAISQGMSVQLHVPYSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRG 1176 1121 QIGAISQGMSVQLHVPYSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRG 1176 1177 QAGPPESLGVPTAQEASVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPAD 1236 1177 QAGPPESLGVPTAQEASVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPAD 1236 1177 QAGPPESLGVPTAQEASVLRGTALGSVPGGSITKGIPSTRVPSERKKGSITHGTPAD 1236 1177 QAGPPESLGVPTAQEASVLRGTALGSVPGGSITKGIPSTRVPSERKKGHVLSYEGGMS 1285 1237 VLYKGTITRIIGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMS 1285 1237 VLYKGTITRIIGEDSPSR
6 6 6 6 6 6 6	3 6 6 6 6 6 6	8 8 8 8 8	6 6 6 6 6 6	6 6 6 6 6 6 6

nidogen.
-!- SUBUNIT: There are six type IV collagen isoforms, alpha 1(IV)alpha 6(IV), each of which can form a triple helix structure with
2 other chains to generate type IV collagen network.
-!- SUBCELLULAR LOCATION: Cell surface (Potential).
-!- TISSUE SPECTRICITY: Alpha 3 and alpha 4 type IV collagens are
collocalized and present only in basement membranes of kidney, eye,
collocalized and brain.
-!- DOMAIN: Alpha chains of type IV collagen have a noncollagenous
domain (NCI) at their C-terminus, frequent interruptions of the GX-Y repeats in the long central triple-helical domain (which may
cause flexibility in the triple helix), and a short N-terminal
proposition of the triple helix. 귱 WEDLINE-99011253; Pubmed-9792860;
BOye E., Mollet G., Forestier L., Cohen-Solal L., Heidet L.,
Cochat P., Gruenfeld J.-P., Palacoux J.-B., Gubler M.-C., Antignac C.;
Determination of the genomic structure of the CoL444 gene and of
novel mutations causing autosomal recessive Alport syndrome.";
Am. J. Hum. Genet. 63:1329-1340(1988).
-I. FUNCTION: Type IV collagen is the major structural component of
glomerular basement membranes (GBM), forming a 'chicken-wire'
meshwork together with laminins, proteoglycans and entactin/ Kamagata Y., Mattei M.-G., Ninomiya Y.,
"Isolation and sequencing of cDNAs and genomic DNAs encoding the
alpha 4 chain of basement membrane collagen type IV and assignment of
the gene to the distal long arm of human chromosome 2.";
J. Biol. Chem. 267:23753-23758(1992). PTM: Prolines at the third position of the tripeptide repeating unit (G-X-Y) are hydroxylated in some or all of the chains.

PTM: Type IV collagens contain numerous cysteine residues which are involved in inter- and intramolecular disulfide bonding. 12 of these, located in the NC1 domain, are conserved in all known type IV collagens.

DISEASE: Defects in COL444 are a cause of autosomal recessive VARIANT AS SER-1201.
MEDLINE=95078927; PubMed=7987396;
Mochizuki T., Lemmink H.H., Mariyama M., Antignac C., Gubler M.-C., Pirson Y., Verellen-Dumoulin C., Chan B., Schroeder C.H., Smeets H.J.M., Reeders S.T.;
"Identification of mutations in the alpha 3(IV) and alpha 4(IV) collagen genes in autosomal recessive Alport syndrome."; Sugimoto M., Oohashi T., Yoshioka H., Matsuo N., Ninomiya Y.; "CDNA isolation and partial gene structure of the human alpha 4(IV) collagen chain."; MEDLINE-96379660; PubMed-8787673; Lemmink H.H., Nillemen W.N., Mochizuki T., Schroeder C.H., Brunner H.G., van Oost B.A., Monnens L.A.H., Smeets H.J.M.; "Benign familial hematuria due to mutation of the type IV collagen alpha4(IV) collagen chains are arranged head-to-head on chromosome MEDLINE=97338662; PubMed=9195222; Lemmink H.H., Schroeder C.H., Monnens L.A.H., Smeets H.J.M.; "The clinical spectrum of type IV collagen mutations."; Hum. Mutat. 9:477-499(1997). Clin. Invest. 98:1114-1118(1996). SEQUENCE OF 1407-1507 FROM N.A. MEDLINE=93054733; PubMed=1429714; MEDLINE=93374047; PubMed=8365481; SEQUENCE OF 1219-1690 FROM N.A. FEBS Lett. 330:122-128(1993). !q36."; 'EBS Lett. 424:11-16(1998). Nat. Genet. 8:77-82(1994). VARIANTS AS, AND VARIANTS. VARIANT FBH GLU-897 REVIEW ON VARIANTS alpha4 gene."; +

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
            characterized by progressive glomerulonephritis, renal failure, hematuria, ocular abnormalities and deafness. The recessive form occurs equally between males and females. The recessive form DISBASE: Defects in COLA44 are a cause of familial benign hematuria (FBH) [MIM:141200] or thin basement membrane disease. FBH is characterized by persistent hematuria, an electron microscopically detectable thin glomerular basement membrane microscopically detectable thin glomerular basement membrane and an autosomal dominant mode of inheritance. Renal function remains normal. In children, differentiation between FBH and AS can be difficult, because both disorders are manifested by SIMILARITY: TO OTHER TYPE IV COLLAGENS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CELL ATTACHMENT SITE (POTENTIAL).
OR 1566 (BY SIMILARITY).
OR 1569 (BY SIMILARITY).
OR 1569 (BY SIMILARITY).
OR 1681 (BY SIMILARITY).
OR 1686 (BY SIMILARITY).
BY SIMILARITY.
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Extracellular matrix, Connective tissue, Basement membrane, Repe
Hydroxylation, Collagen, Glycoprotein, Signal, Disease mutation,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CELL ATTACHMENT SITE (POTENTIAL).
CELL ATTACHMENT SITE (POTENTIAL).
CELL ATTACHMENT SITE (POTENTIAL).
CELL ATTACHMENT SITE (POTENTIAL).
CELL ATTACHMENT SITE (POTENTIAL).
CELL ATTACHMENT SITE (POTENTIAL).
CELL ATTACHMENT SITE (POTENTIAL).
CELL ATTACHMENT SITE (POTENTIAL).
CELL ATTACHMENT SITE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             . .) (POTENTIAL) . .) (POTENTIAL)
 syndrome (AS) [MIM:203780], an hereditary disorder
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COLLAGEN ALPHA 4 (IV) CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /FTId=VAR_008148.
G -> A (in dbSNP:1800516).
/FTId=VAR_008149.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           7S DOMAIN.
TRIPLE-HELICAL REGION.
NONHELICAL REGION (NC1)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /FTId=VAR_008150.
G -> E (in FBH).
/FTId=VAR_001912.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (in AS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Missing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ·
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A -> T.
                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR008161; Clg_helix.
InterPro; IPR008160; Collagen.
InterPro; IPR001442; Procollagn4_C.
Pfam; PF01413; C4; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam; PF01391; Collagen; 21.
ProDom; PD000007; Clg_helix; 3.
ProDom; PD003923; ProcollagnC4; 1.
                                                                                                                                                                                                                                                                                                                          EMBL; X81053; CAA56943.1; -.
EMBL; AB008496; BAA25065.1; -.
EMBL; D17391; BAA0414.1; -.
PIR; A55360; CGHU1B.
Genew; HGNC:2206; COL4A4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    syndrome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            64
1459
1690
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SMART; SM00111; C4; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1214
1569
1566
1531
1686
1683
1683
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              931
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   897
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Polymorphism; Alport
SIGNAL 1 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1525
1588
1622
1634
                                                                                                                                                                                                                                                                                                                                                                                                                    MIM; 141200; -. MIM; 203780; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     570
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              931
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            441
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   897
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SITE
SITE
SITE
SITE
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SITE
```

OY 1515 ASSSGGSIARGAPVIVPELGKRRQSPLTYEDHGAPFAGHLPRGSPVTWREPTPRLQEGSL 1574	0.5 MHNTATAMAQRADMLRGLSPRESSIALNYAAGPRGIIDLSQVPHLPVLVPPTPGTPATA 1754	QY	QY 2046SSYSPEGVE-PVSPVSSPSLTHDKGLPKHLEELDKSHLEGELRPKQPQKLGGGE 2099 Db 1317 CDGKDGQKGPMGFPGPQGPHGFPGPPGEKGLPGPFGRKGPTGLPG-1361 QY 2100 AAHLPHLRPLESQPSSSPLLQTAPGVKGHQRVVTLAQHISEVITQDYTRHHPQQLS 2156 Db 1362PRGEPGPPADVDDCPRIPGLPGAPGMRGPEGAMGLPGMR 1400 QY 2157 APLPAPLYSPPGASCPVLDLRRPPSDLYLPPPDHGAPARGSPHSEGGKRSPEDNKTS 2213 Db 1401 GPPGPGCKGEPGLDGRRGVDGVPGSPGPPGRKGDT 1435	VIGGGEDGIE-PVSPPEGMTEPGHSRSAVY 2242	RESULT 67 TP20_HUMAN STANDARD; PRT; 1960 AA. ID _TR20_HUMAN STANDARD; AC Q9UGÜO; 014528; Q13078; Q9H4M0; DT 28-FEB-2003 (Rel. 41, Created) DT 28-FEB-2003 (Rel. 41, Last sequence update) DT 10-OCT-2003 (Rel. 42, Last annotation update)
FT VARIANT 1004 1004 L-> P (in dbSNP:1800517). FT /FTId=VAR 008151. FT VARIANT 1030 1030 G-> V (in AS). FT VARIANT 1201 1201 G-> S (in AS). FT VARIANT 1402 1402 P-> S (in AS). FT VARIANT 1572 1572 P-> S (in AS). FT VARIANT 1572 1572 P-> L (in AS). FT CONFLICT 1659 1660 P-> L (in AS). FT CONFLICT 1659 1660 A, 164095 MM; EIE72F283A72BAAB CRC64;	Ad. No. 0.00032; Mismatches 617; Indels 681; Gaps 9PPGPPTPPRRTSRAPIEPTPASEAUGAPTPP 80	869 BAEEGPAKGKDAEAAEATAECALKAEKKEGGSGRATTAKSSGAPQDSDSSATCSADEVDE	1035 FAAEAQKLPGDPPCWTSGLPPPV 607 ATPGGKGFPGLGCPGGTGPGGTGGFPG 1085 PLGLHDTARPVLPRPPTISNP 655 PDGLKGQKGDTISCNVTYPGRHGPGFDGP 1136 PYSEHAKAPVGPVTMGLPLPMDDP 103DGHKGRPGTPGTAGIPLPMDDP 703DGHKGRPGTPGTAGIPLPMDD	Qy 1189 AQEASVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYKGTITRI 1246 Db 753 NGQKGIPGDPAFGHLGPPGKRGLSGVPGIKGPRGDPGCP 791 QY 1247 IGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSV 1286 PD 792 -GAEGPAGIPGFLGLKGPKGREGHAGPPGVPGPPGREGRAPGIPGQPGLPGGYPG 846 QY 1287 TOCSKEDGRSSGPPHETAAPKRTYDWMEGRVGRAISSASIBGLMGRAIPPERHSPHHLK 1346 Db 847SPGAPGGKGQPGDVGPP-GPAGWKGLDFGRPGAHGPPCL- 886 QY 1347 EQHHIRGSITQGIPRSYVEAQEDYLRREAKLIKREGTPPPPPPSRDLTEAYKTQALGPLK 1406 ID	OY 1407 L

```
REAL BEDGINGE FROM N. Almostough R. Almeida J.P., Babbage A.K., RA Dunham I., Hunt A.R., Collina J.E., Baresk K.N., Beare D.M., RA Dunham I., Hunt A.R., Collina J.E., Baresk K.N., Beasley O.P., RA Bagguley C., Bailey J., Barlow K.F., Baresk K.N., Beasley O.P., Bailey J., Barlow K.F., Baresk K.N., Beasley O.P., Barlow S.E., Bridgeman A.M., Buck D., Chen Y., Clark G., Corroy D., Corby W.E., Collie G.J., Cox A.V., Davis J., Dawson B., Damis D.D., Dockree C., Doddworth S.J., Durbin R.M., Ellington A.G., Evans K.L., Few J.M., Fleming K., French L., Garner A.A., Hall C., RA Glubert J.G.R., Coward M.E., Graftam D.V., Griffiths M.N.D., Hall C., Rans K.L., Few J.M., Fleming K., French L., Garner A.A., Hall C., Rans K.L., Few J.M., Fleming C., Hall-Tamlyn G., Heathcott R.W., Ho S., Hollmes S., Hunt S.E., Jones M.C., Rexhaw J., Kimberley A.M., King A., Marlesh M.A., Lloyd C., Lloyd D.M., Marley J.M., Malresh M.A., Lloyd C., Lloyd D.M., Marley J.M., Skuce C.D., Smalley S., Morling B.J., Goddil C.N., Pavitt R., Skuce C.D., Smalley S., Small M.L., Ross M.T., Scott C.E., Sharm M.A., Lloyd C., Lloyd D.M., Malliams E., Williams E., Williams B.A., Williams D., Williams E., Williams C., Sharman A., Shibuya K., Saeski T., Asakwa S., Rodoh J., Swan R.M., Milley D.L., Ranger W. Milley D.L., Ranger W. M., Milley D.L., Band M., Milley D.L., Sharman A., Shibuya K., Yochizaki Y., Aoki N., Mituryan S., Do T., Shintan A., Shibuya K., Yochizaki Y., Aoki N., Mituryan S., Do T., Band M., Zhang G., Chen F., Chu P., Malaj E., Wayen T., Pan H., Lowis S., Linds S., Luns S., Williams D., Willer N., Willer T., Scheet P., Walze S., Water D., Scheet P., Walze S., Water S., Williams D., Wares S., Water S., Water S., Water S., Water S., Water S., Water S., Water S., Water S., Water S., Water S., Water S., Water S., Water S., Water S., Water S., Water S., Water S., Water S., Water S., Water S., Water S., Water S., Water S., Water S., Water S., Water S., Water S., Water S., Water S., Water S., Water S., Water S., Water S., Water S., Water S., Water S., 
Transcription factor 20 (Stromelysin 1 PDGF-responsive element-binding
                                                                                                                                                    Rekdal C., Sjoettem E., Johansen T.; The nuclear factor SPBP contains different functional domains and stimulates the activity of various transcriptional activators."; J. Biol. Chem. 275:40288-40300(2000).
                                                   Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Seki N., Nomura N.; "Construction of human brain cDNA libraries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       suitable for analysis of cDNA clones encoding relatively large
                                                                                                            (ISOFORMS 1 AND 2), CHARACTERIZATION, AND
              In) (SPRE-binding protein) (Nuclear factor SPBP) (AR1)
OR SPBP OR KIAA0292.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ohara O., Nagase T., Ishikawa K.-I., Nakajima D., Ohira M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE OF 245-1960 FROM N.A. (ISOFORM 1)
                                                                                                                                        MEDLINE=20568288; PubMed=10995766;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE=97323006; PubMed=9179496;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Nature 402:489-495(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DNA Res. 4:53-59(1997).
                                                                                                           SEQUENCE FROM N.A. (I
                                         Homo sapiens (Human)
                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                          NCBI TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TISSUE=Brain;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       proteins.";
              protein)
TCF20 OR
```

SEQUENCE OF 1083-1938 FROM N.A. (ISOFORM 2)

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the BMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its was by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                   hybridization.";
Cytogenet. Cell Genet. 81:176-177(1998).
Cytogenet. Cell Genet. 81:176-177(1998).
-!- FUNCTION: Transcriptional activator that binds to the regulatory region of MMP3 and thereby controls stromelysin expression. It stimulates the activity of various transcriptional activators such as JUN, SPI, PAX6 and ETS1, suggesting a function as a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TISSUE SPECIFICITY: Expressed in most tissues, except in ovary and prostate. Isoform 1 is exclusively expressed in brain, heart and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE=98399843; PubMed=9730594; Rajadhyaksha A., Riviere M., Van Vooren P., Szpirer J., Szpirer C., Babin J., Bina M.; Habin M.; Hasignment of AR1, transcription factor 20 (TCF20), to human chromosome 22q13.3 with somatic cell hybrids and in situ
                                                                                                                                                                                                                                                                                                 SUBUNIT: Homodimer (Probable). Interacts with RNF4 and JUN (By
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GO: 00:005634; C:nucleus; NAS.
GO; GO:0003713; F:DNA binding; NAS.
GO; GO:0003713; F:transcription co-activator activity; NAS.
GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
InterPro; IPR001965; Znf_PHD.
Transcription regulation; Activator; DNA-binding; Zinc-finger;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LEUCINE TIPPER.

NUCLEAR LOCALIZATION SIGNAL
(BY SIMILARITY).
A.T HOOK.

NUCLEAR LOCALIZATION SIGNAL
(BY SIMILARITY).

NUCLEAR LOCALIZATION SIGNAL
(BY SIMILARITY).
PHD-TYPE (ATYPICAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VRLWR (in isoform 2)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IsoId=Q9UGU0-2; Sequence=VSP 003984, VSP 003985;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PPLPC -> VRLWR (in isof
/FTId=VSP_003984.
Missing (In isoform 2).
/FTId=VSP_003985.
                                                                                                                                                                                                                                                                                                                                                                                                  Event=Alternative splicing; Named isoforms=2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                   IsoId=Q9UGU0-1; Sequence=Displayed;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 protein; Alternative splicing.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POLY-GLN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SER-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POLY-ALA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POLY-GLN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POLY-PRO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; AY007595; ÁAG28930.1; -.
EMBL; AL031346; CAB4240.1; ALT INIT.
EMBL; AL021878; -; NOT ANNOTATED_CDS.
EMBL; ABO06630; BAA22961.1; -.
EMBL; U19345; AAC36392.1; ALT_FRAME.
Genew; HGNC:11631; TCF20.
                                                                                                                                                                                                                                                                                                                                               -!- SUBCELLULAR LOCATION: Nuclear.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 frameshift in position 1932.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1551
1600
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1792
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1960
                                                                                                                                                                                                                                                                               coactivator.
                                                                                                                                                                                                                                                                                                                               similarity).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1537
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             203
310
1556
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MIM; 603107; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1785
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1939
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     884
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 62
174
                                                                                                                                                                                                                                                                                                                                                                                                                              Name=1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Nuclear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ZN FING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DNA
```

COMPLICT COMPLI	1341 1014 1016 1052 1112 1112 1116 11273 1273 1273 1273 1273 1369 1398 1398 1437	OY 1875 PSVLHNTGMGGIITAVEPSKPTVLRS Db 1461 PGAMTSTTSQKP QY 1933 MEPVLLPKEAPRVAREERRADTG 1933 MEPVLLPKEAPRVAREERRADTG 1940 GHATIARTPAKNIAPHHASPDPPAPP QY 1990 GHATIARTPAKNIAPHHASPDPPAPP 1990 GHATIARTPAKNIAPHHASPDPPAPP 1990 GHATIARTPAKNIAPHHASPDPPAPP 1990 GHATIARTPAKNIAPHHASPDPPAPP 1545GSVNKÇKÇQQPFPPPPQPP 1545GSVNKÇKÇQQPFPPPPQPP 1594 RKRKTKQAVPIVEPQEBIKLK QY 2046 LRPKQPGPVKLGGEAAHLPHLRPLPEP 1594 RKRKTKQAVPIVEPQEBIKLK QY 2046 LRPKQPGPVKLGGEAAHLPHLRPLPEP 1594 RKRKTKQAVPIVERQEBIKLK 1650 INAEEBEQTKLVRGRKGQRSLTPPPS 1650 INAEEBEQTKLVRGRKGQRSLTPPPS 110 KWASYRNMGDLFGPFYPQDYAATLP
6 4 6 4 6 4 6 4 6 4 6 4 6 4 6 4 6 4 6 4	SEQUENCE 1960 AA; 211798 MW; FB459146C05B6D2A CRC64; Duery Match 2.64; Score 339; DB 1; Length 1960; Best Local Similarity 19.18; Pred, No. 0.00038; Matches 371; Conservative 231; Mismatches 695; Indels 648; Gaps 407 YRRGGSQQQQQQQQQQ-QQQQPMPRSSQERXDEKEKEAAEKEERPKPUNDKED 5	940 PRPSILTPTGDPRANASPQKPLDLKQLKQRAAAIPPIQVTKVHEPPREDAAPTKPAP-PA 627SQEDDPAATQRP

------PLKLKPAHEGLVATVKEA 1421 .K--EGSITQGTPL--KYDTGASTTGSKKHDVRSLI 1477 KNDKDGMMYGPPVGTYHDPSAQEAG-----RCLM 1165 REP----TPRLOE----GSLSSSKASOD 1582 AFNSYAHL----SH---SQDIKSIPKRDS 1316 HIRGSITQGIPRSYVEAQEDYLRREAKLLKREGTPP 1385 AGLNSQLHYKRQMYQQQPEEYKDWSSGSAQGVIAA 1111 JISPYEHLLRGVSGVDLYRSHIPLAFDPTSIPR--- 1639 YPPYLIRGYPDTAALENROTIINDYITSQQMHHNT 1699 ----GSVDDKVKTETHA------ETVTAGKEP 1460 ---GSNQGRPDGS-----LGGTAPLIFPDSKN 1496 PKHLEELDKSHLEGE----- 2085 AAGPRGIIDLSQVPHLPVLV----PPTPGTPATAM 1755 G-PTHLTKPTTTSSERERDRDRERDRDREREKSI 1814 NOBLHVEKPLPRSSEE------1436 SSGGGGGSSSRPASHSHAHQHSPISPRTQDALQQR 1874 GHAFLAKPPARSGLEPAS-SPSKGSEPRPLVPPVS 1989 P----ASASDPHREKTQSKPFSIQELELRSLGYHG 2045 : |||: : : KYATQPLDKTDAKNKSFYPYIHVVNKCELGAVCTI 1649 ESQ----PSSSPLLQ----TAPGVKGHQRVVTLA 2136 STSTSSPVRPAATFPPATHCPLGGTLDGVYPTL-- 1932 POIPEGSADGEPKPKKÓR-----ÓRRERRKPGAQP 1593 1734

1505 AA

774

831

```
1466 TGSKKHDVRSLIGSPGRTFPPVHPLDVMADARALERACYEESLKSRPGTASSSGGSIARG 1525
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              975 ----PIQVTKVHEPPRED-----AAPTKPAP-PAPPPPQNLQPESDAPQQPGSSP-- 1019
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1020 -----RGKSRSP-----APPADKEAFAAEAQKLPGDPPCWTSGLPFPVPPREVI 1063
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1227 GSITHGTPADVLYKGTITRIIGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSV 1286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1287 TQCSKEDGRSSSGPPHETAAPKRTYDMMEGRVGRAISSASIEGLMGRAIPPERHSPHHLK 1346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1347 EOHHIRGSITOGIPRSYVEAOEDYLRREAKLLKREGTPPPPPPPSRDLTEAYKTQALGPLK 1406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1407 LKPAHEGLVATVKEAGRSIHEIPREELRHTPELPLAPRP-LKEGSITQGTPLKYDTGAST 1465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1526 APVIVPELGKPR-----QSPLTYEDHGAPFAGHLPRGSPVTMREPTPRLQEGSLSSS 1577
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      629 IKMKOF-----LSDEQNVLALRTIQVRQRGSITPRIRTPETGSDDAI--- 670
                                                                                                                      99 -KRSOBABAAFLS---VYKOLIBAPD------PVPVFBAARSL-DDRLQPPSF-- 140
                                                                                                                                                                             ---DPSGQPRDLHTSWKRNPELLSPKEQREG--TSPAGPTLTEGSRLPGIPGKALLTET 195
                                                                                                                                                                                                                            832 AAAPPVEEGEEQKPPAAEE----LAVDTGKAEEPVK------SECTE----EAEEG 873
                                                                                                                                                                                                                                                             196 L----LORNEAEKOKGLOEVOITLAARLGEAEEKIKVIHSALKATOAELLELRRKYDEEA 251
                                                                                                                                                                                                                                                                                                  ----AEAAEATAEGALKAEKKEGGSGRATTAKSSGAPQDSDSSA 919
                                                                                                                                                                                                                                                                                                                                   252 ASKADEVGLIMTNLEKANQRAEAAQREVESLREQLASVNSSIRLACCSPQGPSGDKVNFT 311
                                                                                                                                                                                                                                                                                                                                                                     ------DEVDEAEG---GDKNRLLSPRPSLLTPT 948
                                                                                                                                                                                                                                                                                                                                                                                                                                         949 GD-----974 GD------ 974
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          372 BEKLOAQSDYBEIKTELSILKAMKLASSTCSLPOGMAKPEDSL----LIAKEAFFPTQKF 427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 753 ATASONGAPALVKO-----EEGSGGP--AQAPLPVLSPAAFVOSIIRKVKSEIGD 800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -----GOPNGRAWPRGDEAPVPPEDEAAAGAEDEP------PRTGELKAEGA 876
                                                                                                                                                                                                                                                                                                                                                                                           ------PAFYGAKPPTAP----ATPA-PGPEPLGGPEPADGGGGAAGPGAEEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         QIGAISQGMSVQLHVPYSEHAKAPVGPVTMGL------PLP---MDPKKLAPFSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1167 VKQEQLSPRGQAGPPESLGVPTAQEASVLRGTALGSVPGGSITKGIPSTRVPSDSAITYR
                  NYKKRONLDEILOOHKLKMEKERNARRKKKKAPAAASEEAAFPPVVEDEEMEASGVSGNE
                                         ----ROEESEHSHKHLIELRRE-----FKKNVPEEIREMVA--PVLKSFQAEVVALS---
                                                                                   EEMVEEAEALHASGNEVPRGECSGPATVNNSSDTESIPSPHTEAAKDTGQNGPKPPATLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               428 LLEKPSLLASPEEDPSEDDSIKDSLGTEQSYPSPQQLPPPPGPEDPLSPSPGQPLLGPSL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      488 GPDGTRTFSLSPFPSLASGBRLMMPPA---AFKGEAGGL------LVFP-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1064 KASPHAPDPSAFSYAPPGHPLPLGLHDTARPVLPRPPTISNPPLI---SSAKHPSVLER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AGYFDHHWASDRGLLSRPYASVSP-----SLSS----SLSS----
                                                                                                                                                         775 ADGPPPGPPTPPRRTS--RAPIEPTPASEATGAPTPPPAPP-SPSAPPPVVPKEEKEEET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             571 ÓLDTAEIAFQVKEQLLKHNIGGRVFGHYVLGLSQGSVSEILARPKPWRKLTVKGKEPF--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ------KŚIĽEQAKKE-----IĖSQKĠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 874 PAKGKD------
                                                                                                                                                                                                                                                                                                                                                                     920 TCSA-----
                  555
                                                   23
                                                                                                                                                                                            141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        671
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        801
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             836
                                               셤
                                                                                                                   셤
                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   쉽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ઠે
                                                                                     ò
                                                                                                                                                       ò
                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          83;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             654
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              52
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           595 AITPQQSAELASMELNESSRWTEEEMETAKKGLLEHGRNWSAIARMVGSKTVSQCKNFYF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SIAPKMAANVGSM----FQYWKRFDLRRLQKELNSVASELSA-------
                                                                                                                                                                                             ż
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                factor activity; NAS.
transcription, DNA-dependent; NAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Transcription regulation; Homeobox; DNA-binding; Nuclear protein;
Repeat; Coiled coil.

1 1 COILED COIL (POTENTIAL).

DOMAIN 213 393 CULED COIL (POTENTIAL).

DNA BIND 563 650 CUT 1.
                                                                                     Craniata, Vertebrata; Euteleostomi;
Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                    Res. 4:53-59(1997).
FUNCTION: May be a transcription factor involved in neural
specification. Binds to DNA in a sequence-specific manner (By
                                                                                                                                                                                             Seki
                                                                                                                                                                                                                                                                                                                                                     cDNA libraries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         781;
                                                                                                                                                                                                                                                                                                                                                                   suitable for analysis of cDNA clones encoding relatively large
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 1505;
                                                                                                                                                                                            Ohara O., Nagase T., Ishikawa K.-I., Nakajima D., Ohira M.,
                                                                                                                                                                                                                                                                                                                Ohara O., Nagase T., Ishikawa K.-I., Nakajima D., Ohira M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2.6%; Score 338.5; DB 1; Length :
Larity 19.4%; Pred. No. 0.00031;
Conservative 158; Mismatches 622; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AA; 163792 MW; F68F574B2B3EC41E CRC64;
                                                                                                                                                                                                              Nomura N.;
Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                           similarity).
SUBCELLULAR LOCATION: Nuclear (By similarity).
SIMILARITY: Belongs to the CUT homeobox family.
SIMILARITY: Contains 3 CUT domains.
SIMILARITY: Contains 1 homeobox domain.
                                                                                                                                                                                                                                                                                                                                   Seki N., Nomura N.; "Construction of human brain
28-FEB-2003 (Rel. 41, Last sequence update) 10-0CT-2003 (Rel. 42, Last annotation update) Homeobox protein Cux-2 (Cut-like 2) (Fragment) CUTL2 OR KIAA0293.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COILED COIL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HOMEOBOX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GG 3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE=97323006; PubMed=9179496;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PROSITE; PS00027; HOMEOBOX 1; 1.
PROSITE; PS50071; HOMEOBOX 2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; AB006631; BAA22962.2; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ProDom; PD000010; Homeobox; 1.
SMART; SM00389; HOX; 1.
                                                                                                                                                                                                                                                               SEQUENCE OF 212-1505 FROM N.A.
                                                                                   Eukaryota; Metazoa; Chordata;
Mammalia; Eutheria; Primates;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam; PF02376; CUT; 3.
Pfam; PF00046; homeobox; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             393
650
736
993
1144
                                                                     Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Similarity
                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                      NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1505
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1057
                                                                                                                                                                         ISSUE=Brain;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         375;
                                                                                                                                                                                                                                                                                                                                                                                         proteins."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DNA_BIND
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches
```

g ઠ

```
1757
                                                                                                                                                                                                                                                                        1801 DRERDRDREREKSILTSTTTVEHAPIWRPGTEQSSGSSGSSGGGGGSSSRPASHSHAHQH 1860
                                                                                                                                                                                                                                                                                                                                                                                                                1926
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2068
578 KASQDRKLISTPREIAKSPHSTVPEHHPHPISPYEHLLRGVSGVDLYRSHIPLAFDPTSI 1637
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     969 EPFIRMQLWLSDQLGQAVGQQPGASQASPTEPRSSPSPPPSPTEPEK-----SSQEPLSL 1023
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1227 KTNTVINWFHNYRSRMRREMLVEGTQDEPDLDPSGGPGI-----LPPGHSHPDP--- 1275
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ------GLPKHLEBLDKSHLEGELRPKQ-------PGPVKLGGEA 2100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1319 QEQMEEDAEEEAGSQPQDSGELDK----GQGPPKEEHPDPPGNDGLPKVAPGPLLPGGST 1374
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2101 AHLPHLRPLPESQPSSSPLLQTAPGVKGHQRVVTLAQHISEVITQDYTRHHPQQLS---- 2156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ------RLHPDPLSFKSA 1402
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ----APLPAPLYSFPGASCPVLDLRRPPSDLYLPPPDHGAPARGSPHSEGGKRSPEPNK 2211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1403 SESSRCSLEVSLNSPSAASSPGLMMS-----VSPVPSSSAPISPSPRAPAKVPSASP 1456
                                                                                          1638 PRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAALENROTIINDYITSQQMHH
                                                                                                                                                          1024 SLESSKENQ-----SQAPGGIQEIVAM
                                                                                                                                                                                                                                                                                                                                           1861 SP-----ISPRIODALQ--------QRPSVLHNTGMKGIITA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1174 -----SPCLOPODLSLLOIKKPRVVLAPEEKEALRKAYOLEPYPSOOTIELLSFOINL
                                                                                                                                      1698 NTATAMAQRADMLRGLSPRESSLALNYAAGPRGIIDLSQVPHLPVLVPPTPGTPATAMDR
                                                                                                                                                                                                      1758 LAYL-------PTAPQPFSSRHSSSPLSPGGPTHLTKPTTTSSSERERDR
                                                                                                                                                                                                                                                                                                                                                                                                               -TSTSSPVRPAATFPPATHCPLGGTLD
                                                                                                                                                                                                                                                                                                                                                                                                                                           1123 VRMOLWLNDPHNVEKLRDMKKLEKKAYLKRRYGLISTGSDSESPAT---RSECP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1927 GVYPTLMEPVLLPKE-----APRVARPERPRADTGHAFLAKP-PARSGLEPASSP---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1976 ------SKGSEPRPLVPPVSGHATIARTPAKNLAPHHASPDPPAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2015 PASASDPHREKTQSKPFSIQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDK-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             28-FEB-2003 (Rel. 41, Last sequence update)
10-CCT-2003 (Rel. 42, Last annotation update)
SON protein (SON3) (Negative regulatory element-binding protein) (N
binding protein) (DBP-5) (Bax antagonist selected in saccharomyces
(BASS1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia, Eutheria, Primates, Catarrhini, Hominidae, Homo.
NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SON HUMAN STANDARD; PRT; 2426 AA.
P18583; O14487; O95981; Q14120; Q9H7B1; Q9P070; Q9P072; Q9UKP9;
Q9UPY0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1375 PDCPSLHPQOESEAGE------
                                                                                                                                                                                                                                                                                                                                                                                                           1890 V----EPSKPTVLRS----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2212 TSVLGGGEDGIEPVSP 2227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1457 TADMAGALHPSAKVNP 1472
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-NOV-1990 (Rel. 16, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (Human)
                                 877 TAEAGARL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           28-FEB-2003
10-OCT-2003
                                                                                                     899
                                                                                                                                                                      928
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2069
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2157
                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                    g
                              g
                                                                                               a
                                                                                                                                 ò
                                                                                                                                                                                                    ò
                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                 ò
                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                       ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
```

```
"Transcription repression of human hepatitis B virus genes by negative regulatory element-binding protein/SON."; J. Biol. Chem. 276:24059-24067(2001).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE OF 1-130 FROM N.A.

TISSUB-Smooth muscle;

Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,

Nakajima Y., Mizuno T., Morinaga M., Tanigami A., Fujiwara T., Ono T.,

Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Ota T., Suzuki Y.,

Obayabhi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y.,

Isogai T., Sugano S.;

"NEDO human cDNA sequencing project.";

Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
              SEQUENCE FROM N.A. (ISOFORMS A; B; C; D; E AND F).
MEDLINE=21564202; PubMed=11707072;
Reymond A., Friedli M., Neergaard Henrichsen C., Chapot F.,
Deutsch S., Ucla C., Rossier C., Lyle R., Guipponi M.,
Antonarakis S.E.;
"From PREDs and open reading frames to cDNA isolation: revisiting the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Bliskovskii V.V., Kirillov A.V., Zakhariev V.M., Chumakov I.M.; "The human son gene: the large and small transcripts contains various
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 "Identification of a protein product of a novel human gene SON and the biological effect upon administering a changed form of this gene into mammalian cells.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE-99397452; PubMed=10470851; Kikuno R., Mayajima N., Kikuno R., Nagase T., Ishikawa K.-I., Hirosawa M., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.; "Prediction of the coding sequences of unidentified human genes. XIV The complete sequences of 100 new cDNA clones from brain which code
                                                                                                                                                                                                                                                       C.-K.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TISSUB=Placenta;
MEDLINE=93062884; PubMed=1435773;
Bliskovskii V.V., Berdichevskii F.B., Tkachenko A.V., Belova M.E.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE OF 554-2426 FROM N.A. (ISOFORM A).
MEDLINE=92049296; PubMed=1944255;
Chumakov I.M., Berdichevskii F.B., Sokolova N.V., Reznikov M.V.,
                                                                                                                                                                                                                                                                                                                                                                                                             ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             YE M., Zhang Q.H., Zhou J., Shen Y., Wu X.Y., Guan Z.Q., Wang Fan H.Y., Mao Y.F., Dai M., Huang Q.H., Chen S.J., Chen Z.; "Human partial CDS from cd34+ stem cells."; Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                               TISSUE=Liver;
MEDLINE=21316479; PubMed=11306577;
Sun C.-T., Lo W.-Y., Wang I.-H., Lo Y.-H., Shiou S.-R., Lai
                                                                                                                                                                                                                                                                                                                                                                                                   Casadei R., Strippoli P., D'Addabbo P., Canaider S., Len
Vitale L., Giannone S., Carinci P., Zannotti M.;
Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE OF 437-2426 FROM N.A. (ISOFORM B).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE OF 709-1079 FROM N.A. (ISOFORM I).
                                                                                                                                      human chromosome 21 transcription map.";
                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE OF 1-689 FROM N.A. (ISOFORM H).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Biol. (Mosk) 25:731-740(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Mol. Biol. (Mosk) 26:807-812(1992)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TIŜSUE=Placenta;
MEDLINE=93062885; PubMed=1435774;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  for large proteins in vitro.";
DNA Res. 6:197-205(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE OF 1009-1131 FROM N.A.
                                                                                                                                                                                            SEQUENCE FROM N.A. (ISOFORM G)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE OF 1-114 FROM N.A.
                                                                                                                                                      Genomics 78:46-54(2001).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5'-terminal sequences,
                                                                                                                                                                                                                                                                                                                                                                                       TISSUE-Placenta;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Prasolov V.S.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 rissum=Blood;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TISSUE=Brain;
                                                                                                                                                                                                                                                                         Ting L.-P.;
```

us-09-522-753-5.rsp

```
use by non-profit institutions as long as its content is in inc madified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   67 EVLSGVLDTELRYKPDLK-EGSRK-SRCVSV-----QTDPTDEIPTKKSKKHKKHKKKK 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             742 VNNSSDTESIPSPHTEAAKDTGQNGPKPPATLGADGPPPGPPT--PPRRTSRAPIEPTPA 799
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ** 943 OKPPAAEELAVD--TGKAEEP-----VKSECTEEAEEGPAKGKDAEAAEATAEGALKAEK 895
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         896 KEGGSGRATTAKSSGAPQDSDSSATCSADEVDEAEGGDKORLLSPRPSLLTPTGDPRANA 955
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             173 AFGPSETNESPAVVLB-----PPVVSMEVSEPHILETLKPATKTAELSVVSTSV 221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        682 KKKKAPAAASEEAAFPPVVEDEEMEASGVSGNEEEMVEEAEALHASGNEVPRGECSGPAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 564 VASKGRKTAN--SQGRRKGRITRSMANEANSEEAITPQQSAELASMELNESSRWTEEEME
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    622 TAKKGLLEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDEILQQHKLKMEKERNARR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ----SEATGAPTPPPAPPS-----PSAPPPVVPKEEKEEETAAAPPVE------EGEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  282 TSPEPSKIMLVEPPVAKVLEPSETLVVSSETPTEVYPEPSTSTTMDFPESSAIEALRLPE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    17 X 10 AA TANDEM REPEATS OF L-A-[ST]-
[NSG]-[TS]-MDSQM.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
2.6%; Score 337.5; DB 1; Length 2426;
Best Local Similarity 18.8%; Pred. No. 0.00052;
Matches 457; Conservative 282; Mismatches 899; Indels 789;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         OPVDVPSEIADSSMTRPOELPELPKTTALELOESSVASAMELPGP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GO; GO:0008189; F:apoptosis inhibitor activity; IDA.
GO; GO:0003677; F:DNA binding; TAS.
GO; GO:0006916; P:anti-apoptosis; IDA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SMART; SM00358; DŠRM; 1.
SMART; SM00443; G patch; 1.
PROSITE; PS50137; DS.RBD; 1.
PROSITE; PS50174; G FATCH; 1.
RNA-binding; DNA-binding; Nuclear protein; Repeat;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL, AB028942; BA482971.1; --
EMBL, X63071; CA444793.1; ALT_FRAME.
EMBL, AF138997; AAD50078.1; --
EMBL, AK024752; BAB14985.1; --
EMBL, AF161428; AAF28988.1; --
EMBL, AF161430; AAF28988.1; --
EMBL, AF161430; AAF28988.1; --
EMBL, AF161430; AAF28988.1; --
Genew, HGNC:11183; SON.
GK; P18583; --
MIM, 182465; --
                                                                                                                                                                  EMBL, AF380179; AAL34497.1; -.
EMBL, X63753; CAA45282.1; ALT_FRAME
EMBL, M36428; AAA36624.1; -.
                                                                                                                                                                                                                                                              EMBL, AF380180, AAL34498.1, EMBL, AF380181, AAL34499.1, EMBL, AF380182, AAL34500.1, EMBL, AF380183, AAL34501.1, EMBL, AF380184, AAL34502.1, EMBL, AY026895, AAK07692.1,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; AF435977; AAL30810.1; -. EMBL; X63751; CAC69885.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR001159; DS RBD.
InterPro; IPR000467; G patch.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam; PF00035; dsrm; 1.
Pfam; PF01585; G-patch; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Alternative splicing DOMAIN 726 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            342
THE REAL PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE OF THE PRINCE 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SFRS2/SC-15.
SIMILARITY: Contains 1 G-patch domain.
SIMILARITY: Contains 1 DRBW (double-stranded RNA-binding) domain.
CAUTION: Isoform A sequence from Ref.7 differs from that shown due to a frameshift.
CAUTION: Isoform F sequence from Ref.10 differs from that shown due to a frameshift.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      IsoId=P18583-10; Sequence=VSP_004414, VSP_004415; TISSUE SPECIFICITY: Widely expressed, with the higher expression seen in leukocyte and heart.

DOMAIN: Contains 8 types of repeats which are distributed in 3
                            Coding part of the son gene small transcript contains four areas of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE=89039788; PubMed=3054499;
Berdichevskii F.B., Chumakov I.M., Kiselev L.L.;
Berdichevskii F.B., Chumakov I.M., Kiselev L.L.;
"Decoding of the primary structure of the son3 region in human peromen; identification of a new protein with unusual structure and homology with DNA-binding proteins.";
Mol. Biol. (Mosk) 22:794-801(1988)."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE OF 1939-2426 FROM N.A. (ISOFORM J).
TISSUB-Cerebellum;
MEDLINE-99439804, PubMed=10509013;
Greenhalf W., Lee J., Chaudhuri B.;
"A selection system for human apoptosis inhibitors using yeast.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IsoId=P18583-2; Sequence=VSP_004401, VSP_004402, VSP_004403;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                regions. MISCELLANEOUS: Colocalizes with the pre-mRNA splicing factor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Event=Alternative splicing; Named isoforms=10; Comment=Experimental confirmation may be lacking for some
                                                                                                                                                                                                                                     2
                                                                                                                                                           SEQUENCE OF 1145-2426 FROM N.A. (ISOFORM P).
MEDLINE-93048367; PubMed-1424986;
Mattioni T., Hume C.R., Konigorski S., Hayes P., Osterweil
                                                                                                                                                                                                                                                                  Lee J.S.; "A cDNA clone for a novel nuclear protein with DNA binding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 VSP_004407;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VSP_004412;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VSP_004405;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    VSP_004409
                                                                                                                                                                                                                                                                                                                                                                                                                                            (ISOFORM A).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IsoId=P18583-3; Sequence=VSP_004404,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Name=B;
IsoId=P18583-6; Sequence=VSP_004408,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IsoId=P18583-8; Sequence=VSP_004411,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IsoId=P18583-4; Sequence=VSP_004406,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IsoId=P18583-7; Sequence=VSP_004410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      IsoId=P18583-9; Sequence=VSP_004413;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 IsoId=P18583-5; Sequence=VSP_004403;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    IsoId=P18583-1; Sequence=Displayed;
                                                               ete tandem repeats.";
Biol. (Mosk) 26:793-806(1992)
                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE OF 1692-2175 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                        Chromosoma 101:618-624(1992).
                                                            tandem
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        soforms;
```

Name=A;

Name=C

<u>:</u>

104;

Gaps

621

681

99

842

281

--- 386

Q9952-1 USULES; QUESTED CONTROL OF CHEER CONTROL OF CHEER CONTROL OF CHEER CONTROL OF CHEER CONTROL OF CHEER CONTROL OF CHEER CONTROL OF CHEER CONTROL OF CHEER CONTROL OF CHEER CONTROL OF CONTROL OF CHEER CONTROL OF CONTROL OF CHEER CONTROL OF CHEER CONTROL OF CONTROL OF CONTROL OF CHEER CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CO	F L L L L L L L L L L L L L L L L L L L	1236 DPSVLVSEAAVTVPEPPPEPESSITLTPVESAVVAEEH 1273 1859 QHSPISPRTQDALQQRPSVLHNTGMKGIITAVEPSKPTVLRSTS 1902	da ya
MUSB_HUMAN ID MUSB_HUMAN STANDARD; PRT; 570	RES MUS ID	1750 TPATAMDRLAYLPTAPQPPSSRHSSSPLSPGGPTHLTKPTTTSSSERERDR 1800	ර් සි
2494		1733PG 1749 1119 TADRSMMSMAADSYTDGYTDTYTEAXMVPPLPPEBPPTMPPLPPEBPPWTPPLPPEBPPE 1178	ÇO Q
LTDLDKAQLLEIAKANAAAMCAKAGVPI	අ _ධ :	1732 1059 MSAYERSMMSAYERSMMSPWADRSMMSMGADRSMMSSYSAADRSMMSSYSAADRSMMSSYI118	2 8
1993 RTPSRRRRSRSVVRRRSFSISPVRLR 2462PPYNPLIMRLQAGVMAS	8 & ——	RSM	면 전 -
24—t	ර් ස්		· &
	qa	955 PYRLTPDPYRMSPRPYRIAPRSYRIAPRPYRIAPRPYRIAPRPYRIAPPRPYRIAPPRYRIA	ි <u>සි</u> .
1875 KSKSKGRRSVSKEKRRKSPKHRSKSRE 2379 AAMPITAADGRSDHTLTSPGGGGKAKV	89	901 QDSAMLGSKSPDPYRLAQDPYRLAQDPYRLGHDPYRLGHDAYRLGQDFYRLGHD 954	ଶ (
	\dot \d	1570 QEGSLSSSKASQDRKLTSTPREIAKSPHSTVPEHHPPISPYEHLLRGVSGVDLYR-SHI 1628	ò
RSKRS	qa —	1514 TASSEGGESPYTMYEPTEL 1569 1514 TASSEGGESPYTMYEPTEL 1569 1515 T	3 8
1755 PLLASDVGRDRSAASPVVSSMPERASESS	Q	814 QMLATSSMDSQMLATSSMDSQMLATSTMDSQMLATSTWDSQMLATS 857	qq
	λο Ολ	PGRTFPPVHPLDVMADARALERACYEESLKSRPG	ò
	qa	 MLATSSMDS	; A
1048SGGSEADIEGFLEFAKDIHLDLE 2210 NKTSVLGGGEDGIEPVSPPEGMTEPGHSR	₹ <i>\delta</i>	719 NIMETHILASNIMDSQMLASNIMDSQMLASNIMDSQMLASSIMDSQMLASSMDSQ 774	<u> </u>
	<i>₹</i>	EDYLRREAKULKREGTPPPPPPSRDL-TEAYKTO	ò
1592 TLSSTGPFALEPDATGTSKGIEFTTAS		1290 SKEDGRSSSGPPHETAAPKRIYDMMEGRVGRAISSASIEGLMGRAIPPERHSPHH 1344 1290 SKEDGRSSSGPPHETAAPKRIYDMMEGRVGRAISSASIEGLMGRAIPPERHSPHH 1344 665 STWIVSQSLEVPSITALESTVIVAQELPTILVGETSYTYGFDFMAPERHILAS 718	ờ A
NHLIAKEMEHNTVCAAG	du ,	613 GAAGALELLGQPLATGVLELPGQPGAPELPGQPVATVALEISVQSVVTTSEL	q
2093		1230 THGTPADVLYKGTITRIIGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQC 12	ò
2042 GYHGSSYSPEGVEPVSP-VSSPSLTHDKC : : : 1483 VMKGINLS-SGDQNLAPEIGMQEIALHSG	ਨੇ ਜ਼ਿ 	1177 QAGPPESLGVPTAQEASVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSI 1229	& 8 8
			7 A
1381 VIVLEPSVVIVPEPPVVAEPD		468 EVPEPSVMAQELPCLPLVTAAVELPEQPAUTVAMELTEQPVTTTELEQ	<u>ል</u> :
	λο :	1061 EVIKASPHAPDPSAFSYAPPGHPLPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLER	ò
	qa —	434 AVPELPGP	음·
		387PATSMPELQGPPVTPVLELPGPSATPVPELPGPLSTPVPELPGPPAT 1005 LOPKSDAPOODGSGPPGKGBSBABBABABABART DANDATHINGGT DED	8 8
		956 SPQKPLDLKQLKQRAAAIPPIQVTKVHEPPREDAAPTKPAPPAPPPQN 1004	ò

|:| |: |: ------ASEPPVMSETAETFDSMRASGH 1320 KGLPKHLEELDKSHLEGELRPKQPG----- 2092 ATHCPLGG-----TLDGVY 1929 AMAAPESSAMAVLESSAVTVLESSTVTVLESST 1380 PAPPASASD --- PHREKTQSKPFSIOELELRSL 2041 AHLPHLRPLPESQPSSSPLLQTAPGV---- 2126 : || |: -----KILPTSETKQRTVLDTYPGVSEADAGE 1591 ACHIS-----EVITODYTRHHPQQLSAPL 2159 PPDHG------APARGSPHSEGGKRSPEP 2209 | :| | :| ASTLSLVNKYDVDLSLTTQD-TEHDMVISTSP- 1647 SRS-----AVYPLLYRDGEQTEPSRMG--- 2258 ------KOEI 2290 EIFNMPAITGTGLMTYRSQAVQEHASTNMGLEA 2344 | : ::||| : ::||| SBRSRARRES 1874 ------PLSANAFNPLNASAS-LP 2378 CVSGRPS----SRKAKSPA-----PGLASGD 2425 OCNRRIPLINRV------WEDRPSSAGSTP-- 2461 -RSRTPLRRRFSRSPIRRKRSRSSERGRSPKR 2049 SP---- PPPGLPAGSGPLAG----- 2493 | | | | : |PPNLKPAPPPTIEEKVAKKSGGATIEELTEKC 2109 ;703 AA. 14; 095291; 095451; Q14881; Last sequence update)
 Last annotation update)
 (Mucin 5 subtype B, tracheobronchial) (High 2136

```
SEQUENCE OF 1-1594 FROM N.A.
Chen Y., Di Y.P., Wu R.;
"Molecular cloning of the amino-terminal and 5'-flanking region of the
human WUCSB mucin gene.";
                                                                                                                                                                                                                                                                                      MEDLINE=99023932; PubMed=9804771;
Desseyn J.-L., Buisine M.P., Porchet N., Aubert J.-P., Laine A.;
Menomic organization of the human mucin gene MUCSB: cDNA and genomic sequences upstream of the large central exon.";
J. Biol. Chem. 273:30157-30164(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 'Identification of a major human high molecular weight salivary mucin (MG1) as tracheobronchial mucin MUC5B.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                "Genomic organization of the 3 region of the human MUC5B mucin.";
"Genomic organization of the 3 region of the human MUC5B mucin.";
J. Biol. Chem. 272:16873-16883(1997).
-!- FUNCTION: Salivary mucin that is thought to contribute to the lubricating and viscoelastic properties of whole saliva.
-!- SUBCELLULAR LOCATION: Secreted.
-!- TISSUE SPECIFICITY: Expressed mainly in bronchus glands and also
                                                                                                                                                                                                                                                                                                                                                                                                                                  "Human mucin gene WUC5B, the 10.7 kb large central exon encodes various alternate subdomains resulting in a super-repeat. Structural evidence for a 11p15.5 gene family.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Keates A.C., Nunes D.P., Afdhal N.H., Troxler R.F., Offner G.D.; "Molecular cloning of a major human gall bladder mucin: complete C-terminal sequence and genomic organization of MUC5B."; Biochem. J. 324:295-303(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Iroxler R.F., Offner G.D., Zhang F., Iontcheva I., Oppenheim F.G.;
Molecular cloning of a novel high molecular weight mucin (MG1)
from human sublingual gland.";
                                                                                                                                                                       SEQUENCE OF 1-1325 FROM N.A.
MEDLINE=99009274; PubMed=9790959;
Offiner G.D., Nunes D.P., Keates A.C., Afdhal N.H., Troxler R.F.;
"The amino-terminal sequence of MUC5B contains conserved
multifunctional D domains: implications for tissue-specific mucin
                                Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
molecular weight salivary mucin MG1) (Sublingual gland mucin).
MUCSB OR MUCS.
                                                                                                                                                                                                                                                                                                                                                                                                              Desseyn J.-L., Guyonnet-Duperat V., Porchet N., Aubert J.-P.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TISSUE-Salivary gland;
BDLINE-97292540; PubMed-9147051;
Nielsen P.A., Bennett E.P., Wandall H.H., Therkildsen M.H.,
Hannibal J., Clausen H.;
                                                                                                                                                Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Biochem. Biophys. Res. Commun. 217:1112-1119(1995)
                                                                                                                                                                                                                                                     Biochem. Biophys. Res. Commun. 251:350-355(1998)
                                                                                                                                                                                                                                                                                                                                                                                   TISSUE=Placenta;
MEDLINE=97166151; PubMed=9013550;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TISSUE=Sublingual gland;
MEDLINE=96125355; PubMed=8554565;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE=97347489; PubMed=9201995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE=97293229; PubMed=9164870;
                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE OF 1326-4895 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE OF 4057-4480 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE OF 4721-5703 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE OF 4809-5687 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE OF 4859-5703 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Glycobiology 7:413-419(1997).
                                                                                                                                                                                                                                                                              SEQUENCE OF 40-1324 FROM N.A.
                              Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 rissum=Gall bladder;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FISSUE=Placenta;
                                                                                                                                                                                                                                          functions.
                                                                                                                                                                                                                                                                                                                                                                                                                           Laine A.;
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
              PTM: Highly glycosylated.
SIMILARITY: Contains 1 TIL (Trypsin inhibitory-like) domain.
SIMILARITY: Contains 3 VMFC domains.
SIMILARITY: Contains 4 VMFD domains.
SIMILARITY: Contains 1 C-terminal cystine knot-like (CTCK) domain.
in submaxillary glands, endocervix, gall bladder, and pancreas.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ....
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CTCK.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Polymorphism.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VWFD 2.
VWFC 1.
VWFD 3.
THR-RICH.
THR-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MIM, 600770; -.
GO; GO:0005515; F:protein binding; IPI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MUCIN SB.
VWFD 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VWFD 4.
VWFC 2.
VWFC 3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR006208; Cya knot.
InterPro; IPR006207; Cya knot. C.
InterPro; IPR009041; PMP_inhibitor.
InterPro; IPR002919; VIL_Cyarich.
InterPro; IPR001007; VWF_C.
                                                                                                                                                                                                                                                EMBL, AF107890; AAG33673.1; --
EMBL, AF086604; AAC67545.1; --
EMBL, AJ004862; CAA06167.1; --
EMBL, Z72496; CAA06167.1; --
EMBL, X74955; CAA0530.1; --
EMBL, U7855; CAA52910.1; --
EMBL, U7855; AAC51344.1; JOINED.
EMBL, U7855; AAC51344.1; JOINED.
EMBL, U7855; AAC51343.1; --
EMBL, U95031; AAC51343.1; --
EMBL, Y09788; CAA70926.1; --
GGDDGW, HGNC;7516; MUCSB.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR001007; VWF_C.
InterPro; IPR001007; VWF_C.
Pfam; PP00007; Cys_knot; 1.
Pfam; PF00093; VWc; 1.
Pfam; PF00094; VWc; 4.
SMART; SM00214; VWC; 6.
SMART; SM00215; VWC out; 4.
PROSITE; PS01185; CTCK 1; 1.
PROSITE; PS01185; CTCK 2; 1.
PROSITE; PS01125; CTCK 2; 1.
PROSITE; PS01125; CTCK 2; 1.
PROSITE; PS01125; CTCK 2; 1.
PROSITE; PS01126; VWFC_1; 2.
PROSITE; PS01126; CTCK 2; 1.
PROSITE; PS01126; CTCK 2; 1.
PROSITE; PS01126; CTCK 2; 1.
PROSITE; PS01126; CTCK 2; 1.
PROSITE; PS01126; CTCK 2; 1.
PROSITE; PS01126; CTCK 2; 1.
PROSITE; PS01126; CTCK 2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1044
1603
4873
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5178
5178
5178
5178
5178
5178
5178
5182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
```

TYWHEPPREDAPPTRAPPAPPONIQPESDAPOOPGSSPRGKSRSPAPPADKEAPAAE TTATTP - TVTSKATPSSSPGTATALP - ALESTATPTATTTPATTTPATTTPATTTPATTTPATTTPATT	1866 RDREREKSILTSTTTVEHAPIWRPGTEQSSGSSGSSGGGGGSSSRPASHSHAHQHSPI 1863 3156 STRATAGTLKVLTSTATTPTVISSRATPSS-SPGTATALPALRSTA 3200 1864 SPRTQDALQQRPSVLHNTGMKGIITAVEPSKPTVLRSTSTSSPVRPAATFPPATHCPLGG 1923 1201 TTPTATSVTAIPSSSLGTAWTRLSQTTTPTATMKSTATPSST-PETVHTSTVL 3251 1924 TLDGVYPTLMEPVLLPKEAPRVARPERPRADTGHAFLAKPPARRGLEPASSP 1975
3	6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6
930 930 N-LINKED GLCNAC (FOTENTIAL)	EEKEBETAAAPPVEGGEGKPPAAELAVDTGKAEEPVKSECTEEAEGPAKG 877
N-LINKED (GLCNAC. N-LINKED (GL	PPAAEEL
HTML 1277 N-LINKED (GLCNAC) (POTENTIAL)	
FT CARBOHYD FT CONFLICT FT CON	Oy 825 Db 2312 Oy 878 Db 2372 Oy 919

```
2288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2251
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3686
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2394
3252 TTTAT-TTRTGSVATPSSTPGTAHTTKVPTTTTTGFTATPSSSPGTALTPPVWISTTTTP 3310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                S---FTAIPSSSLGTTWTR-----LSQTTTPMATMSTATPSSTPETV-----HTSTVL 3779
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3780 TT----TATTTGATGSVATPSSTPGTAHTTKVPTTTTGFTVTPSSSPGTARTP---PV 3831
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3832 WISTTITPITSGSTVIPSSIPGITHTPIVLITITQPVATGSMAIPSSSIQISGIPP 3887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2450 W---EDRPSSAGSTPFF----YNPLIMR-----LQAGVMASPPPFGLPAGSGP 2490
                                                                                                                                                                                                                                                                   |: | | : : | | : : | | 3366 ----TITATGSTINPSSTPGTIPIPPULITITAATISSTVIPS--SALGITH----
                                                                         1976 -SKGSEPRPLVPPVSGH-ATIARTPAKNLAPHHASPDPPAPPASASDPHREKTQSKPFSI
                                                                                                                                       3311 TTRGSTVTPSSIPGTTHTATVLTTTTTVATGSMATPSSSTQTSGTPPSLTTTAT----
                                                                                                                                                                                                                        2034 OELELRSLGYHGSSYSPE---GVEPVSPV----SSPSLTHDKGLPKHLEELDKSHLEGE
                                                                                                                                                                                                                                                                                                                                                                       2086 LRPKQPGPVKLGGEAAHLPHLRPLPESQPSSSPLLQTA--PGVKGHQRVVTLAQHISEVI
                                                                                                                                                                                                                                                                                                                                                                                                                                             3414 -TPPVP-----NTTATTHGRSLPPSSPHTVPTAWTSATSGILG-----TTHITE--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2144 TQDYTRHHPQQLSAPLPAPLYSFPGASCPVLDLRRPPSDLYLPPPDHGAPARGSPHSEGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----PSTGTSHTPAATTGTTQPSTPALSSPHPSSRTTESPPSPGTTTPG--HTRGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ------GGGED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SRTTATATPSKTRTSTLLPSSPTSAPITTVVTTGCEPQCAWSEWLDYSYPMPGPSGGDFD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -GIE---PVSPPEGMTEPGH--SRSAVYPLLYRDGEQ-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3567 TYSNIRAAGGAVCEOPLGLECRAOAOPGVPLRELGOVVECŚLDFGLVCRNREÓVGKFRMC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ------ESNSAMVKSKKQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3627 FNYEIRVFCCNYGHCPSTPATSSTATPSSTPGTTWILIKLTTTATTTESTGSTATPSSTÖ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ------EINKKLNTHNRNEPEYNISQPGTEIFNMPAITGTGLMTYRSQAVQEHAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3687 GPPAGTPHVSTTATTPTVTSSKATP---FSSPG------TATALPALRSTATTPTAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TNMGLEAIIRKALMGKYDQWEESPPLSANAFNPLNASASLPAAMPITAADGRSDHT---L
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TSPGGGGKAKVSGRPSSRKAKSPAPGLASGDRPPSVSS----VHSEGDCNRRTPLTNRV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Saarela J., Ylikarppa R., Rehn M., Purmonen S., Pihlajaniemi T.; "Complete primary structure of two variant forms of human type XVIII collagen and tissue-specific differences in the expression of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE=20289799; PubMed=10830953;
Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Groner Y.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Collagen alpha 1(XVIII) chain precursor [Contains: Endostatin]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           P39060; Q9UK38; Q9X6Q7; Q9X6Q8;
01-FRB1-1995 (Rel. 31, Created)
0-CT-2001 (Rel. 40, Last sequence update)
15-MAR-2004 (Rel. 43, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1516 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      KR----SPEPNKTSVL------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE=98164096; PubMed=9503365;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   corresponding transcripts.";
Matrix Biol. 16:319-328(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CA1H HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2338
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3735
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COL18A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CAIH HUMAN
ID CAIH HUMAN
ID CAIH H
DT 01-FEBO
DT 16-OCT
DT 15-MAR
DE COLIBA
COL BARATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
OC BURATY
                                                                     ઠે
                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                    Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
```

```
Soeda E., Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D., Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W., Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S., Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E., Minoshima S., Shimizu, Nordasek G., Hornischer K., Brandt P., Schoen O., Desario A., Reichelt J., Kauer G., Bloecker H., Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E., Wehrmeyer S., Borzym K., Cardiner K., Nizetic D., Francis F., Lehzach H., Reinhardt R., Yaspo M.-L.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Iughetti P., Suzuki O., Godoi P.H., Alves V.A., Sertie A.L.,
Zorick T., Soares F., Camargo A., Moreira E.S., di Loreto C.,
Moreira-Filho C.A., Simpson A., Oliva G., Passos-Bueno M.R.;
"A polymorphism in endostatin, an angiogenesis inhibitor, predisposes
for the development of prostatic adenocarcinoma.";
Cancer Res. 61:7375-7378(2001).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 "Collagen XVIII, containing an endogenous inhibitor of angiogenesis and tumor growth, plays a critical role in the maintenance of retinal structure and in neural tube closure."; Hum. Mol. Genet. 9:2051-2058(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FUNCTION: COLA18A probably plays a major role in determining the retinal structure as well as in the closure of the neural tube. FUNCTION: Endostatin potently inhibits endothelial cell proliferation and angiogenesis. May inhibit angiogenesis by binding to the heparan sulphate protecoglycans involved in growth factor signalling.
                                                                                                                                                                                                                                       SEQUENCE OF 834-1516 FROM N.A. MEDINE-94245237; PubMed=8188291; On S.P., Warman M.L., Seldin M.F., Cheng S., Knoll J.H., Timmons S., Olsen B.R.; Containg of containing of containing of containing of containing the alpha 1KVVIII) collagen and localization of the alpha 1KVVIII) collagen gene to mouse chromosome 10 and human chromosome 21.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      detachment, macular abnormalities and occipital encephalocele. SIMILARITY: BELONGS TO THE FIBRIL-ASSOCIATED COLLAGENS WITH INTERRUPTED HELICES (FACIT) FAMILY.
SIMILARITY: Concains 1 TSP N-terminal (TSPN) domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                      Zhi-Yong H., Biao L., Wei-Jie Z., Xiang-Pu W.; "Cloning and expression of human endostatin gene in Escherichia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   .
:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE=20400145; PubMed=10942434;
Sertie A.L., Sossi V., Camargo A.A., Zatz M., Brahe
Passos-Bueno M.R.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Name=Long; Synonyms=NC-493;
IsoId=P39060-1; Sequence=Displayed;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE=21518361; PubMed=11606364;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 INVOLVEMENT IN KNOBLOCH SYNDROME
                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE OF 1334-1516 FROM N.A.
                                                                                                                                                                                                                                                                                                                                              chromosome 10 and human chi
Genomics 19:494-499(1994).
                                                                                                                                                                                                       Nature 405:311-319(2000)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VARIANT ASN-1437.
                                                                                                                                                                                                                                                                                                                                                                                                                     TISSUE=Placenta;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             <u>:</u>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              <del>!</del>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                                                     [4]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2
```

```
1064
1084
1120
1123
1126
1206
1304
1314
1324
1443
                                                                                                                                                                           886
912
933
933
1064
11084
11123
11126
11206
11314
11314
11443
11483
11516
                                                                                                                                                                                                                                                CONFLICT
                                                                                                                                                                                                                                                                                                     CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        598
                                                                                                                                                                                                                                 CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  953
                                                                  VARIANT
             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          \stackrel{\triangleright}{\sigma}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the Buropaen Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                              CG): GG: 0005581; C: collagen; TAS.

RG): GG: 0005581; C: collagen; TAS.

RG): GG: 0005581; P: histogenesis and organogenesis; TAS.

GG: 00108281; P: histogenesis and organogenesis; TAS.

GG: 00108281; P: histogenesis and organogenesis; TAS.

RG): GG: 0007601; P: vision; TAS.

RG): GG: 0007601; P: vision; TAS.

RICEPPO: IPR008160; Collagen.

RICEPPO: IPR008185; Cond_like_lec_gl.

RICEPPO: IPR008185; Cond_like_lec_gl.

RICEPPO: IPR0081819; Laminin_G.

RICEPPO: IPR0081819; TSPN.

RETOROM: PF01311; Collagen; 7.

RETOROM: PF01311; Collagen; 7.

RETOROM: PF01301; Collagen; 7.

RETARCELLUAR MART; SM00210; TSPN; 1.

RETARCELLUAR MART; SM00210; TSPN; 1.

WEXTERCELLUAR MART; Connective tissue; Repeat; Hydroxylation; Cell adhesion; Collagen; Glycoprotein; Signal; Alternative splicing; SIGNAL

TENTAL SIGNAL

TENTAL SIGNAL

TENTAL SIGNAL

TENTAL SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BY SINILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY GELL ATTACHMENT SITE (POTENTIAL).

Missing (in isoform short).

/FIIdavSp 001155.

HTTEAGTLPAPTPSPPSLGRPWAPLTGPSVPPPSS -> MA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TERP N-TERMINAL.

NOWHELICAL REGION 1 (NC1).

TRIPLE-HELICAL REGION 1 (COL1).

NOWHELICAL REGION 2 (COL2).

TRIPLE-HELICAL REGION 3 (NC2).

TRIPLE-HELICAL REGION 3 (NC2).

TRIPLE-HELICAL REGION 4 (NC4).

TRIPLE-HELICAL REGION 4 (COL4).

NOWHELICAL REGION 5 (NC5).

TRIPLE-HELICAL REGION 6 (COL5).

NOWHELICAL REGION 6 (COL5).

TRIPLE-HELICAL REGION 7 (COL5).

NOWHELICAL REGION 9 (COL6).

NOWHELICAL REGION 9 (COL6).

TRIPLE-HELICAL REGION 7 (COL7).

TRIPLE-HELICAL REGION 9 (COL9).

NOWHELICAL REGION 9 (NC9).

TRIPLE-HELICAL REGION 10 (NC1).

TRIPLE-HELICAL REGION 10 (NC1).

TRIPLE-HELICAL REGION 10 (NC1).

TRIPLE-HELICAL REGION 10 (NC1).

TRIPLE-HELICAL REGION 10 (NC10).

TRIPLE-HELICAL REGION 10 (NC10).

TRIPLE-HELICAL REGION 10 (COL10).

TRIPLE-HELICAL REGION 10 (COL10).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          . .) (POTENTIAL)
. .) (POTENTIAL)
. .) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED (GLCNAC. ...) (POTENTIAL)
N-LINKED (GLCNAC. ...) (POTENTIAL)
N-LINKED (GLCNAC. ...) (POTENTIAL)
N-LINKED (GLCNAC. ...) (POTENTIAL)
O-LINKED (GALNAC. ...) (FOTENTIAL)
/FIId=CAR_000150.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COLLAGEN ALPHA 1 (XVIII) CHAIN ENDOSTATIN.
                                                                                                                                                          EMBL, AF018081, AAC39658.1;
EMBL, AF018082, AAC39659.1;
EMBL, AL163302, CABS0482.1;
EMBL, L22548; AA51864.1;
EMBL, AF184060, AAF01310.1; ALT_INIT.
PDB; 1BNL, 02-DEC-98.
GlycoSuiteDB; P39060; -.
Genew; HGNC:2195; COL18A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1516
5106
5109
5109
5109
5109
5109
6640
6640
6640
6640
6640
10034
10034
10034
10034
10034
10034
10034
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
10039
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1498
1097
180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1366
1468
1095
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SITE
VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
CHAIN
DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
```

1012 PQQPGS-----DXE----- SPRGKSRSPAPPA------DXE---- 1033 1034 -AFAAEAQK-----LPGDPPCWTSGLP------FP-VPPREVIKASPHAPD 1071 1132 QLHVPYSEHAKAPVG-PVTMGLPLPMDPKKLAPFSGVKQEQLS--PRGQAGPPESLGVPT 1188 964 KOLKORAAAIP-PIQVTKVHEPPRE--DAAPTKPAPPAP----PP-----QNLQPESDA 1011 .072 PSAFSYAPPGHPLPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSV 1131 1189 AQEASVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYKGTITRIIG 1248 734 492 686 GRFGVNSSDVPGPAGLPGV--PGREGPPGFPGLPGPPGREGPPGRTGQKGSLGEAGA 743 848 AEELAVDIGKAEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAK 907 908 SSGAPQDSDSS----ATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRANASPQKPLDL 963 630 PQGPPGPPGPSFRHDKLTFIDMEGSGFGGDLEALRGPR-GFPGPPGPPGVPGLPGEP--- 685 903 -LSVP-----GPEGRPGFAGFPGPAGPKGNL---GSKGERGSPGPKGEKGEPGSIFSPD 952 675 KERNARRKKKKAPAASBEAAFPPVVEDEEMEASGVSGNEEEMVEEAEALHASGNEVPRG REETGAALKPRLPAP-----PPVTTPPLAGGSSTEDSRSEEVERQTTVASLGAQT---735 ECSGPATVNNSSDTESIPSPHTEAAKDTGQNG----PKPPATLGADGPP--PGPPTRR TSRAPIEPT-PASEATGAPTPPPAPPSPSAPPPVVPKEEKEEFTAAAPPVEEGEEQKPPA 493 -LPGSDSVSTWDGSVRTPGGRVKEGGLKGOKGEPGVPGPPGRAGPPGSPCLPGPPGLPCP 552 VS--PLGPAGPALQTVPGPQGPPGRDGTP-----GRDGE----PGDPGEDGKPG-744 PGHKGSKGAPGPAGARGESGLAGAPGPAGPPGPPGPPGPPGPPGPPGPPDDMEGSGGPFWS ; Score 335; DB 1; Length 1516; ; Pred. No. 0.0004; 89; Mismatches 519; Indels 420; Gaps 861 GDRGSRGEKGDP---GKDGVGQPGLPGPP--GPPGPVV-----YVSEQDGSV----------YKGEI----PRCPWPWPRRRELLDVLAPLVLLLGVRAASAEP (in of developing 1443 R -> T (IN REF. 4) 1483 S -> Y (IN REF. 4). AA; 153840 MW; 3C70F29A4476EE76 CRC64; risk/FIId=VAR 012709.
F -> S (IN REF 2).
I -> V (IN REF 2).
V -> L (IN REF 3).
P -> R (IN REF 3).
P -> R (IN REF 3).
P -> L (IN REF 3).
P -> L (IN REF 3).
P -> L (IN REF 3).
P -> L (IN REF 3).
P -> L (IN REF 3).
P -> L (IN REF 3).
P -> A (IN REF 3).
P -> A (IN REF 3).
P -> A (IN REF 3).
P -> A (IN REF 3).
P -> A (IN REF 3).
P -> A (IN REF 3).
P -> PPGP (IN REF 3).
P -> PPGP (IN REF 3).
R -> G (IN REF 3).
A -> G (IN REF 3).
C -> C (IN REF 3).
D -> A (IN REF 3). /FTId=VSP 001156. D -> N (increased prostate cancer) FTId=VSP 001 SHLARG 1 2.5%; Similarity 22.8%; Best Local Similarity 22.8 Matches 304; Conservative

us-09-522-753-5.rsp

```
Curr. Biol. 12:1052-1058(2002).
[2]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1NW3; 25-MAR-03
607375; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PDB;
      1455
                                                                                                                                                                                                                                                                                                                                                                                                                                               1163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1572
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1682
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1733
                                                                                                                                                                                                                                                                                                                            1123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PLKYDTGASTTGSKKHDVRSLIGSPGRTFPPVHPLDVMADARALERACYEESLKSRPGTA 1515
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1573 SLSSSKASQDRKLTSTPREIAKSPHSTVPEH------HPHPISPYEHLLRGVSGVDL 1623
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----RLTESYCETWRTEAPSATGOASSLLGGRL 1490
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETAAPK
                                                                                                                                                                                                                                                                                                   1071 PGPPGQFPFDFLQLEAEMKGEKGDRGDAGQKGERGEPGGGGFFGSSLPGPPGP-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1734 LSQVPHLPVLVPPTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPG-----GPTHLT
                                                                                        986 -GFPGRPGRPGMNGL-----KGEKGEPGDASLGFGM-----RGMPGPPGPPGPP-
                                                                                                                                                 RTYDMMEGRVGRAISSASI-----EGLMGRAIPPERHSPHHLKEQHHIRGSITQGI
                                                                                                                                                                                                     -----GPPGTPVYDSNVFAESSRPGPPGLPGNQGPP---GPKGAK-----GEV--GP
                                                                                                                                                                                                                                                                  ----TPPPPPSRDLTEA
                                                                                                                                                                                                                                                                                                                                                                                     1397 YKTQALGPLKLKPAHEGLVATVKEAGRSIHEIPREELRHTPELPLAPRPL-KEGSITQGT
                                                                                                                                                                                                                                                                                                                                                                                                                                               ------PGPRGYPG1PGPKGESIRGQP-----GPPGPQGPPGIGYEG--RQGP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1516 SSSGGSIARGAPVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSPVTMREPTP---RLQEG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1202 GASSG-----VRLWATRQAMIGQ-----VHEVPEGWLIFVAEQEELYVRVQNG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 :: | |:| | : | | |:| | |:| | |:| | |:| | |:| | |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1624 YRSHIPLAFDPISIPRGIPLDAAAAYYLPRHLAP-NPIYPHLYPPYLIRGYPDTAALENR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LAS-----PPRLPEPQPYPGAPHHSSYVHLRPARPTSP---PAHSHRDFQPVLHLVAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1683 QTIIN-----DYITSQOMHHNTATAMAQRADMLRGLSPRESSLALNYAAGPRGIID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1349 NSPLSGGMRGIRGADFQCFQQ----ARAVGLAGTFRAFLSSRLQDL---YSIVRRA--D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1787 KPTTTSSSERERDRDRERDRDREREKSILTSTTTVEHAPIWRPGTEQSSGSSGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               疹
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A. (ISOFORM 1), ENZYMATIC ACTIVITY, AND MUTAGENESIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Zhang Y.;
"Methylation of H3-lysine 79 is mediated by a new family of HWTases
without a SBT domain.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Feng Q., Wang H., Ng H.H., Erdjument-Bromage H., Tempst P., Struhl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Craniata; Vertebrata; Buteleostomi; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOTL HUMAN STANDARD; PRT; 1739 AA.

OBTEK3; O60379; Q96JL1;
28-FRB-2003 (Rel. 41, Created)
28-FRB-2003 (Rel. 43, Last annotation update)
15-MAR-2004 (Rel. 43, Last annotation update)
15-MAR-2019 (Rel. 43, Last annotation update)
16-MAR-2019 (Rel. 43, Last annotation update)
16-MAR-2019 (Rel. 43, Last annotation update)
16-MAR-2019 (Rel. 43, Last annotation update)
17-MAR-2019 (Rel. 43, Last annotation update)
18-MAR-2019 (Rel. 43, Last annotation update)
19-MAR-2019 (Rel. 43, L
                                                                                                                                                                                                                                                                  PRSYVEAQEDYLRREAKLLKREG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            163-GLY--GLY-165.
MEDLINE-22119601; PubMed=12123582;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Bukaryota; Metazoa; Chordata;
Mammalia; Butheria; Primates;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1444 HPTWPQKSVWHGSDPNGR-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1847 SSSRPASHSHAH 1858
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1491 LGQSAASCHHAY 1502
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=9606;
                                                                                                                                                 1309
                                                                                                                                                                                                                                                                  1360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1299
                                 1249
                                                                                                                                                                                                     1029
                                                                                                                                                                                                                                                                                                                                                                                                                                               1124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       'L_HUMAN
                                                                                                                                              δ
                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                               Š
                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 쉱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
SEQUENCE OF 43-936 FROM N.A.
Lamerdin J.B., McCready P.M., Skowronski B., Adamson A.W.,
Burkhart-Schultz K., Gordon L., Kyle A., Ramirez M., Stilwagen S.,
Phan H., Velasco N., Garnes J., Danganan L., Poundstone P.,
Christensen M., Georgescu A., Avila J., Liu S., Attix C., Andreise T.,
Trankheim M., Amico-Kraller G., Coefield J., Duarte S., Lucas S.,
Bruce R., Thomas P., Quan G., Kronmiller B., Arallano A.,
Montgomery M., Ow D., Nolan M., Trong S., Kobayashi A., Olsen A.O.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       "Characterization of long cDNA clones from human adult spleen. II. The complete sequences of 81 cDNA clones."; DNA Res. 10:49-57(2003).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FUNCTION: Histone methyltransferase. Methylates Lys-79 of histone H3. Nucleosomes are preferred as substrate compared to free histones.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CATALYTIC ACTIVITY: S-adenosyl-L-methionine + histone L-lysine = S-adenosyl-L-homocysteine + histone N(6)-methyl-L-lysine.
SUBCELLUIAR LOCATION: Nuclear (Probable).
ALTERNATIVE PRODUCTS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          the existence of another mechanism for methylation of lysine residues of histones. CAUTION: Ref.2 sequence differs from that shown due to erroneous
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.; "Prediction of the coding sequences of unidentified human genes. XX. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro."; DNA Res. 8:85-95(2001).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 1-351.
MEDLINE=22515528; PubMed=12628190;
Min J., Feng Q., Li Z., Zhang Y., Xu R.-M.;
Min Structure of the catalytic domain of human DOTIL, a non-SET domain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                suggesting
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TISSUE-Spleen;
MEDLINE-22579292; PubMed-12693554;
Jikuya H., Takano J., Kikuno R., Hirosawa M., Nagase T., Nomura N.,
                                                                                                                                                                                                                                                                                                                                                                                        "Sequence analysis of a 3.5 Mb contig in human 19p13.3 containing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Isold=O8TEK3-2; Sequence=VSP_002228, VSP_002229; Note=No experimental confirmation available; MISCELLANBOUS: In contrast to other lysine histone methyltransferase, it does not contain a SET domain,
                                                                                                                                                                                                                                                                                                                                                                                                                                   serine protease gene cluster.";
Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Event=Alternative splicing; Named isoforms=2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE OF 454-1739 FROM N.A. (ISOFORM 2)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         nucleosomal histone methyltransferase.";
Cell 112:711-723(2003).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IsoId=Q8TEK3-1; Sequence=Displayed;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PARTIAL SEQUENCE FROM N.A. (ISOFORM 1)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TISSUE=Brain;
MEDLINE=21245130; PubMed=11347906;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              gene model prediction.
                                                                                                                                                                                                                                                                                                                                                  Carrano A.V.;
```

983 EPPREDAAPTKPAPPSSPR 1020 673LRGKGALGRELEPDASRILHLELDCTKFSLPHLSSMSPELSWNGQAAGYELCGULSRPS 730 1021 GKSRSPAPPADKZAPAARAQKLPGDPPCWTSGLPFPVPRRUIKASPHAPDDSAFSY 1077 731 SKQNTPQYLASPLDQEVVPCTPSHVGRPRLEKLSGLAAPDYTRLSP 776 1078 APPGHPLPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHVPY 1137 777 AKIVLRRHLSQDHTVPGRPAASELHSRAEHTKENGLPY 814 1138 SEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESL 1184 815QSPSVPGSWLGSPAASELHSRAEHTKENGLPY 849 1185 GVPTAQBASVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHG 1232 870 PLSTVQPNKLPVSIPLASVVLPSRAERARSTPSPVLQPRDPSSTLEKQIGANAHGAGSRS 929 1233TPADVLYKGTITRIIGEDSGREDSLPKGHV 1268 930 LALAPAGESYAGSVAISGALAGSPASLTPGAEEATLDESSSGSLFATVGSRSSTPQHPL 989	1269 1
8 6 8 6 8 6 8 6 8 6	6 B 6 B 6 B 6 B 6 B 6 B 6 B 6 B 6 B 6 B
Transferase; Methyltransferase; Chromatin regulator; Nuclear protein; KW Transferase; Methyltransferase; Chromatin regulator; Nuclear protein; KW Repeat; Alternative splicing; 3D-structure; Polymorphism.	2.5%; Scor 2, Conservative 194; M VLYYLITKURENYKSIVERSYERR

8 8 셤

ð

요 ò

8

```
Similarity
                                                                                                                                                                                                                                     AB024071;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         713
                                                                                                                                                                                                                                                                    AB024073;
                                                AB024060;
                                                                                                                                                   AB024066;
                                                                                                                                                                                  AB024069;
                                                                                                                                                                                                    AB024068;
                                                                                                                                                                                                                   AB024070;
                                                                                                                                                                                                                                                   AB024072;
                                  AB024059
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches 356;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            728
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   950
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MUTAGEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                   Genew;
MIM; 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                 EMBL;
EMBL;
                                                                                                 EMBL;
                                                                                                                                                                 EMBL;
EMBL;
                                                EMBL;
                                                                                                                                                                                                    EMBL;
                                                                                                                                                                                                                   EMBL;
                                                                                                                                                                                                                                     EMBL;
                                                                                                                                                                                                                                                                                    EMBL;
 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   . 염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
-----EPRPLVPPVSGHATIARTPAKNLAPHHASPD----PPAPPASASDPHREKTQ 2027
                                                                     SVETRPPPPPPPPPPPPPAHLGRSPAGPPVLHAPPPPNAALPPPPTLLASNPEPALLQ 1646
                                                                                                                SKPPSIQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDKGLPKHLEELDKSHLEGELR 2087
                                                                                                                                                                                  2088 PKQPGPVKLGGEAAHLPHLRPLPESQPSSSP----LLQTAPGVKGHQRVVTLAQHISEV 2142
                                                                                                                                                                                                                1667 --PPANASLSIKUTSLPHKGARPSFTVHHQPLPRLALAQAAPGIPQASATGPSAVWVSLG 1724
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE=98094256; PubMed=9434167;
Takeuchi T., Chen B.-K., Qiu Y., Sonobe H., Ohtsuki Y.;
"Molecular cloning and expression of a novel human cDNA containing CAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      "The human SWI-SNF complex protein p270 is an ARID family member with non-sequence-specific DNA binding activity."; Mol. Cell. Biol. 20:3137-3146(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SUBCELLULAR LOCATION: Nuclear (Potential)
SIBSUE SPECIFICITY: Highly expressed in aplean, thymus, prostate, testis, ovary, small intestine, colon, and pbl, and at a much lower level in heart, brain, placenta, lung, liver, skeletal smucle, kidney, and pancreas.
SIMILARITY: Contains 1 ARID domain.
CAUTION: Ref. 2 sequence differs from that shown due to a frameshift in position 1132.
                                                                                                                                                                                                                                                                                                                                                                SMF1 HUMAN STANDARD, PRT; 1902 AA.
014497; Q9UPZ1;
16-0CT-2001 (Rel. 40, Created)
16-0CT-2001 (Rel. 40, Last sequence update)
15-MAR-2004 (Rel. 43, Last annotation update)
SWI/SNR-related, matrix-associated, actin-dependent regulator of chromatin subfamily F member 1 (SWI-SNF complex protein p270) (B120).
SMARCPI OR CIORF4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A., AND MUTAGENESIS.
MEDLINE=20221560; PubMed=10757798;
Dallas P.B., Pacchione S., Wilsker D., Bowrin V., Kobayashi R.,
                                                                                                                                                s-----Lasi-----ppnqaflpptsaasi-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE OF 1-1132 FROM N.A.
Takeuchi T., Misaki A.;
Submitted (FEB-1599) to the EMBL/GenBank/DDBJ databases.
-!- FUNCTION: BINDS DNA NON-SPECIFICALLY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; AF265208; AAF75765.1; ALT INIT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE OF 1-1175 FROM N.A.
                                                                                                                                                                                                                                                   2143 ITQDYTRH 2150
                                                                                                                                                                                                                                                                                  1725 MPPPYAAH 1732
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              repeats.";
Gene 204:71-77(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Moran E.;
                                                                                                                2028
                                                                                                                                                1647
```

```
GNEVPRGECSGPATVNNS----SDTESIPSPHTEAAKDTGONGPKPPATLGADGPP 779
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DPRANASPOKPLDLKOLKORAAAIPPIQVTKVHEPPREDAAP-----TKP----AP 996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SQQSTTQQHPQSQPPXSQPQAQSPXQQQQPQQPAPSTLSQQAAYPQPQSQQSQQTAYSQQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  780 PGPPTPPRRTSRAPIEPTPASEATGAPTPPPAPPS-PSAPPPV--VPKEEKEEETAAAPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PQQQQPP--YSQQPPSQTPHAQPSYQQQPQSQPPQLQSSQPPYSQQPSQPPAPQSPAPYP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                837 VEEGEEQ-----KPPAAEELAVDIGKAEEPVKSECIEEAEEGPAKGKDAEAAEATAEGAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                892 KAEKKEGGSGRATTAKSSGAPQDSDSSATCSADEVDEAEGG--DKNRLLSPRPSLLTPTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NGOYATSGWVRDVS (IN REF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GSQTPQ---RYPMTVQGRAQSAMGGLSYTQQIPPYGQQGPSGYGQQGQTP--YYNQQSPH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                       W->A: PARTIAL LOSS OF DNA-BINDING
ACTIVITY. WHEN A.A. 713 ALSO MUTATED,
COMPLETE LOSS OF ACTIVITY.
Y->A: PARTIAL LOSS OF DNA-BINDING
ACTIVITY. WHEN A.A. 690 ALSO MUTATED,
COMPLETE LOSS OF ACTIVITY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels 609;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 1902;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  9B70A46F81062EAA CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ilarity 19.4%; Pred. No. 0.00059;
Conservative 202; Mismatches 664;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 332.5;
                                                                                                                                                                                                                                                                                                                                                                                         POLY-GLN.
POLY-SER.
ARID.
GLN-RICH.
                                                                                            JOINED.
JOINED.
JOINED.
JOINED.
                                                              JOINED.
                                                                                                                                                                           JOINED.
                                                                                                                                                                                                                                                                        GO; GO:0003677; F:DNA binding; NAS.
InterPro; IPR001606; ARID.
InterPro; IPR008938; ARM.
                               JOINED
                                                JOINED
                                                                                                                                             JOINED
                                                                                                                                                               JOINED
                                                                                                                                                                                                                                                            GO; GO:0005634; C:nucleus; NAS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  634 634
924 924
1016 1016
1033 1033
1902 AA; 205946 M
                                                                                                                                                                                                                                                                                                                                         SMART; SM00501; BRIGHT; 1.
DNA-binding; Nuclear protein.
DOMAIN 96 99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2.5%;
   BAA83073.1;
                BAA83073.1;
                               BAA83073.1;
                                                                                          BAA83073.1;
                                                                                                           BAA83073.1;
BAA83073.1;
                                                                                                                                                                                            AB024073; BAA83073.1;
AB024074; BAA83073.1;
                                                              BAA83073.1;
BAA83073.1;
                                                                                                                                               BAA83073.1;
                                                                                                                                                            BAA83073.1
                                                BAA83073.1
                                                                                                                                                                                                                            HGNC:11110; SMARCF1
                                                                                                                                                                                                                                                                                                                          Pfam; PF01388; ARID; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         713
```

949

891

94

Db 1143VNRTDEMLE	1919	1190 QP	OY 1978GSEPRPLVE : Db 1222 PFLHGGMKMQKAGF	Qy 2028 SKPFSIQELELRSI	Db 1268 PVLKORRRITMKDI	2068	1322		1382	Db 1442 VDCSDKIGRYOEPD	Z178 RPPSDLYLPPDHG	1492	Qy 2233 EPGHSRSAVYPLLY	Db 1530 DMLSTRSSTLT	RESULT 74	SON MOUSE STANDARD	AC 090X47; 09C012; 09C0K6; DT 28-FEB-2003 (Rel. 41, C	D1 28-FEB-2003 (Rel. 41, DT 28-FEB-2003 (Rel. 41, DF COM PED-2003	DE SON PIOCETII. GN SON. OS Mis miscrilis (Maisce)	Eukaryota; Metazoa;	OC NCBL TaxID=10090; RN [1]	0, 0,		Zammit P., Dadrah "Organization and	RT and human genomes."; RL Genomics 68:57-62(2000)		RC STRAIN=C57BL/6J; TISSUB: RX MEDLINE=2108560; Pubme	RA Arakawa T., Hara A., Ful		Fleischmann Kuehl P.	RA Schriml L.M., Staubli F RA Sakai K., Okido T., Furn	RA Blake J., Boffelli D., RA Brownstein M.I Bult
: : : 257 -PDLSGSIDDLPMGTEGALSPGVSTSGISSSQGEGSNPAGSPFSPHLPGIRG 311	997 PAPPPDKEAFAABAQX 1042		1043 PGDPPCWTSGLP-PPVPPREVIKASPHAPDPSAFSYAPPGHPL-PLGLHDTAR 1093		: : :	1141 AKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEASVLRG 1197	478 ASMGNRPYGPNMANMPPQVGSGMCPPPGGMNRKT-QETAVAMH 519	1198 TALGSVPGGSITKGIPSTRVPSDSALTYRGSITHGTPADVLYKGTITRIIGEDSPSRLDR 1257	520 VAANSIQNRPPGYPNNNQGGMMGTGPPYGQGINSMAGMINPQ 561		GPPYSMGGTMANNSAGMAASPEMMGLGDVKLTPATKMNNKADGTPKTESKSKKSSSST	1309KEQH 1349 1309KEQH 1349 1309KEQH 1349 620 TINEKITKIYELGREPERRAMINIPYIA PERRAMAMANIANI DANGGERI OF VOLVEY	HIRESTION		1386PPPPSRDLTEAYKTQALGPLKLKPAHEGLVATVKEAGRSIHEIP 1429	733 DIFAAADSKKSQPKIQPPSPAGSGSMQGPQTPQSTSSSMAEGGDLKPPTP 782	1430 REBLRHTPELPLAPRPLKEGSITQGTPLKYDTGASTTGSKKHDVRSLIGSPGRTFPPVHP 1489	-S	LDVMADARA :	835 SDMMGRMSYEPNKDPYGSMRKAPGSDPFMSSGQGPNGGMGDPYSRAAGPGLGNV 888	PRQSPLTYEDHGAPFA	AMGPRQHYPYGGPYDRVRTEPGIGPE	IAKSPYEHLLRGV	GMISFSKIFFUQUQUQUQUQAHIBSYGNQFSTQGTFSSGSPFPSQQTTWYQQQQQNYKRPMDGT 993	HLYPPYLIRGYPDTAA	GENDOMITINDVITEOONHUMMANANAOBANIA DOLLA SERVICE SERVIC		HLPVLVPPTPGTPATAMDRLAYLPTAPODPSSBHSSSPLSPCTFTHTTTSCTPTPT	FQFGRDRVSAPPGTNAQQNMPPQMMGGP	GGGSSSRPASHAH		1859 QHSPISPRIQDALQQRPSVLHNTGMKGIITAVEPSKPTVLRSTSTSSPVRPAATFPPAIH 1918
qa	ò	Ωp	ð 6	ò	. ⁶	ò	дG	ð.	ф	à i	g d	දු ද	8	g Q	ò	qa	ò	qq	ò	g	ð í	g	λo d	3 8	Š	3 8	² 점	ò	q	ò	qa	ò

2099 LHTDQRANHEGSWPSHGTRQPPYGP-----SAPVPPMTRPPPSNY 1189 TLMEPVLLPKEAPRVARPERPRADTGHAFLAKPPARSGLEPASSPSK- 1977 VPPVSGHATIARTPAKNLAPHHASP----DPPAPPASASDPHREKTQ 2027 SLGYHGSSYSPEGVEPVSPVSSPSLTHD-------- 2067 SEEEVVENDEELAFSGKDKPASENSEEKLISKFDKLPVKIVQKNDPFV 1441 DIG-----TPEAMRVMMSLKSGLLAESTWALDTINILLYDDNSIMTF 1321 ------KGHQRV 2132 ------ITQDYTRHHPQQLSAPLPAPLYSFPGASCPVLDLR 2177 PDSGLLHWRIGGGDTTEHIQTHFESKTELLP----SRPHAPC----- 1491 IGAPARGSPHSEGGKRSPEPNKTSVLGGGEDGIEPVSPPE----GMT 2232 IVTTÄEGTPGTT--------DQEGPPPDGPPEKRITATMD 1529 When the state of D-----KOPGPVKLGGE led=10950926;
., Pagel C., Price M., Liu Q.Y., Khan I.M.,
Mazani W., Kessling A., Lee J.S., Buluwela L.;
servation of the GART/SON/DONSON locus in mouse ordata; Craniata; Vertebrata; Euteleostomi; dentia; Sciurognathi; Muridae; Murinae; Mus. TB=Hippocampus, Small intestine, and Tongue; [ed=11217851; YRDG-EQTEPSRMGSKSP 2262 Last sequence update) Last annotation update) PRT; 2404 AA. OFORMS 1 AND 2). RD; K6; Q9QXP5;

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
Mynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
Hayashizaki Y.; Fordida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
T. "Punctional annotation of a full-length mouse cDNA collection.";
I. Nature 409:685-690(2001).
I. PUNCTION: Transcriptional repressor. Binds to the consensus DNA sequence: 5'-da(GT)AN(CG)[AdG]CC-3'. Might protect cells from apoptosis. Might be involved in pre-mRNA splicing (By similarity).
C. -!- SUBCELLUAR LOCATION:
C. -!- ALTERNATIVE PRODUCTS:
Event_Alternative splicing; Named isoforms=2;
                                                                                                                                                                                                                                                                                                 -1- SIMILARITY: Contains 1 G-patch domain.
-!- SIMILARITY: Contains 1 DRBM (double-stranded RNA-binding) domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         13 X 10 AA TANDEM REPEATS OF L-A-[ST]-
[NSG]-[TS]-MDSQN.

11 X 7 AA TANDEM REPEATS OF [DR]-P-Y-R-
11 [AS] [(DHP].

14 X 6 AA REPEATS OF [ED]-R-S-M-M-S.
                                                                                                                                                                                                                                            Isold=090X47-2; Sequence=VSP_004416, VSP_004417;
TISSUE SPECIFICITY: Widely expressed.
DOMAIN: Contains 8 types of repeats which are distributed in 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GO; GO: 0005515; P:protein binding; IPI.
InterPro; IPR00165; DS_RBD.
InterPro; IPR000467; G_patch.
Pfam; PF01585; G-patch; 1.
SMART; SM00443; G_patch; 1.
PROSITE; PS50174; DS_RBD; 1.
PROSITE; PS50174; G_PATCH; 1.
RNA-binding; DNA-binding; Nuclear protein; Repeat;
Alternative splicing.
DOMAIN
721 850
                                                                                                                                                                                                                IsoId=Q9QX47-1; Sequence=Displayed;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AF193595, AAF23120.1; JOINED. AF193595, AAF23120.1; JOINED. AF193597, AAF23120.1; JOINED. AF193599; AAF23120.1; JOINED. AF193601, AAF23120.1; JOINED. AF193601, AAF23120.1; JOINED. AAF93602; AAF23120.1; JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 11-11.
11-21.
11-43.
11-6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; AF193606; AAF23120.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BAB25691.1; -. BAB25562.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAF23121.1;
BAB31659.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            BAB31536.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          943
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 966
974
981
990
998
1006
1015
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; AK008478; BAB25
EMBL; AK008256; BAB25
MGD; MGI:98353; Son.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AF193605; AF193607; AF193607; AK019312; E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                961
969
976
985
993
1001
1010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AF193604;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AX019081;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AF193603;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          867
                                                                                                                                                                                                                                                                                          regions
                                                                                                                                                                                                      Name=1;
                                                                                                                                                                                                                               Name=2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REPEAT
REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL;
EMBL;
EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL;
```

		òòo	666 104 718 159	764 218 821 260	873 315 932 369	991 408 1044 450 1096 499
11-9. 1-10. 1-11. 1-12. 1-13. 1-14. 3 X 11 AA TANDEM REPATS OF P-L-P-P-E-E-P-P-F-E-E-P-P-F-E-E-P-P-F-E-E-F-F-F-F	2-5. 2-6. 2-7. 2-7. 2-8. 2-7. 2-8. 2-8. 2-8. 2-8. 2-9. 2-8. 2-9. 2-8. 2-9. 3-1. 3-1. 3-1. 3-2. 3-1. 3-2. 3-1. 3-2. 3-3. 3-3. 3-3. 3-3. 3-3. 3-3. 3-3	B AS	MELNESSRWTEBEMETAKKGILEHGRNWSAIARWYGSKTVSQCKNFYFNYKKRQNLDBIL	LVLEPPVVSMEVQESHVLETL-KPATKDTG LVLEPPVVSMEVQESHVLETL-KPATKAAELSVVSTSV RRTSRAPIEPTPASEATGAPTPPPAAPSPSAPPPV RRTSRAPIEPTPASEATGAPTPPPAAPSPSAPPPV : : :	LAVDTGKAEEPVKSECTEEAEEG ::	LKQRAAAIPPIQVTKVHEPPREDAAP
1026 1031 1035 1040 1044 1049 1055 1060 1075 1080 1101 1133 1910 1979 1910 1979 1938 1944 1945 1951	959 1965 966 1972 913 1979 919 1937 980 1990 991 2017 283 2329 349 2404 086 2086 087 2404	Similarity 19.2%; P. Similarity 19.2%; P. S. Conservative 260; VASKGRKTANSQGRRKGRI	MELNESSRWTEBEMETAKK 	EEAEALHASGNEVPRGECSGPA	VPKEEKEEETAAAPPVEEGEEQKPPAAEEL	DKORLLSPRPSLLTPTGDPRANASPQKP
REPEAT REPEAT REPEAT REPEAT REPEAT REPEAT DOMAIN REPEAT REPEAT REPEAT REPEAT REPEAT	REPEAT REPEAT REPEAT DOMAIN REPEAT DOMAIN DOMAIN VARSPLIC VARSPLIC SEQUENCE	ery Match St Local Eches 41 564	607 71 667 105	719 160 765 219	822 261 874 316	933 370 . ' 992 409 1045
	8		8 8 8 8	& a & a	8 8 8 8	8 8 8 8 8

<i>ର</i> ଅ	8 A	හි <u>කි</u>	<i>ර</i> සි	8 H	<i>6</i> 8	8 à	i 6	D 1	<i>&</i> 5	RE	255		88888	2	RR	388	3888	8888	88888
1097PRPPIISNPPPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEHAKAPVGPVTWGL 1152	1153 PLPMDPKKLAPPSGVKQEQLSPRGQAGPPESLGVPTAQEASVLRGTALGSVPG 1205	GSITKGIPSTRVPSDSALTYRGSITHGTPADVLYKGTITRIIGEDSPSRLDRGREDSLPK	1266 GHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETAAPKRTYDMMEGRVGRAISSA 1325	SIEGLMGRAIPPERHSPHHLKEQHHIRGSITQGIPRSYVEAQEDYLRREAKLLKREGTPP	PPPPSRDLTEAYKTQALGPLKUKPAHEGLVATVKEAGRSIHEIPREEL-RHTPELPLAPR	00% QELFILLUGGEISVIVGVDFLMAQESHMLASNIMEIHMLASNIMDSQMLASNIMDSQMLAS 743 1445 PLKEGSITQGTPLKYDTGASTTGSKKHDVRSLIGSPGRTFPPVHPLDVWADARALE 1500 : : : : : :	NTMDSOMLASSTWDSQMLASSTWDSQMLATSTWDSQMLATSSMDSQMLA	1501 RACYEESLKSRPGTASSSGGSIARGAPVIVPELGKPRQSPLTYEDHGAPPAGH 1553	1554 LPRGSPVTWREPTPRLQEGSLSSSKASQDRKLTSTPREIAKSPHSTVPEHHPHPISPYEH 1613 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	1614 LLRGVSGVDLYR-SHIPLAFDPTSI-PRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYL 1669			ALNYAAGPRGII	DRLAYLPTAPQPPSSRHSSSPLSPGGPTHLTKPTTTSSSERERDRDREREKSIL	1111BQSAL 1143 1816 TSTTTVEHAPIWRPGTEQSSGSSGSGGGSSSRPASHSHAHQHSPISPRTQDALQQRP 1875	1144 TADNTWSTEVTLSTGESLSQPEPPVSQSEISEPMAV-PANYSMSESET 1190	1876 SVLHNTGMKGIITAVEPSKPTVLRSTSTSSPVRPAATPPPATHCPLGGTLDGV 1928 	1929 YPTLMEPVLLPKEAPRVARPERPRADTGHAFLAKPPARSGLB 1970	
\$ A	& g	상 원	8 8	8 8	3 8 8	8 8	<u>염</u>	දු පු	95 15 15 15 15 15 15 15 15 15 15 15 15 15	රු සි	8 8	상 원	& g	ठे ह	8 8	g	දු අ	ço q	ζ, 19

```
CA12_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ξ
                                         g
                                                                                                                                                 ď
                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                    쉽
                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                            ଚ
                                                                                                                                                                                                                                                                                                                                                                                                                                ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             مير
                      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMED outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              31;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | | | | : | | : | | : | | : | | : | | : | | : | | : | | : | | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | : | | : | | : | : | | : | : | | : | : | | : | : | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           503
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      551
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             611
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        KOMROLAVIPPMLYDADOORIKFINMNGL---MADPMKVYKDROVMNMWSEOEKETFREK 443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               731
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -------ELRGTKWEMARHLREYODLLNVKMALDIEIAAYRKL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           444 FMQHPKNFGLIASFLERKTVAECVLYYYLTKKNENYKSLVRRSYRRGKSQQQQQQQQQQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   504 QOQQQQPMPRSSQEEXDEXEXEXEREEXPEVE-----NDKEDLLKEKTD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BLIQNMDRVDREITMV-----EQQISKLKKKQQQLEBEAAKPPEPEKPVSPPIESKH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        224 RSLVQIIYDENRKKAEAAHRILEGLGPQVELPLYNQPSDTRQYHENIK------INQA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  M---RKKLILYFKRRNHARKQWKOKFCORYDQLMEALEKKVERIENNPRRAKESKVREY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          333 YEK-----OFPEIRKORELQERMOSRVGQRGSGLSMSAARSEHEVSEIIDGLSEQENLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LEGEETRFSTFAG-----QPSIAISS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DISGEDNDEKEAVASKGRKTANSQGRRKGRITRSMANEANSEEAITPQQSAELASMELNE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ----KSPVKATAPELKEEEGEKEEEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SSRWTEEEMETAKKGLLEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDEILQQHKL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               KMEKERNARRKKKKKAPAASEEAAFPPVVEDEEMEASGVSGNEEEMVEEAEALHASGNEV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           252;
                                                                                                                                                                                                                                                                                                                                                     Length 810;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative 119; Mismatches 307; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LINKER 2.
COIL 2B.
8 X 13 AA TANDEM REPEATS.
B8477D85560AC3F6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ------GS-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2.5%; Score 332; DB 1;
21.2%; Pred. No. 0.00028;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TAIL.
COIL 1A.
LINKER 1.
COIL 1B.
LINKER 12.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               : :: |:| ||:|
EKEEKEEAEBEVVAAK------
                                                                                                                                                                                                                                                              Interpro; IPR001664; IP.
Interpro; IPR002957; Keratin I.
Pfam; PF00038; filament; I.
PRINTS; PR01248; TYPELKERATIN.
PROSITE; PS00226; IF; I.
                                                                                                                                                                                                                                  EMBL; AF091342; AAC36357.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 90799 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GOEEEEEEEAAKSDOAEEG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    132
149
171
175
296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               <1
21
34
133
150
176
176
503
810 AA;</pre>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches 182;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     612
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               276
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    387
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             552
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
SO SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITION OF SET TITITIO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               요
```

```
PRGECSGPATVNN-----SSDTESIPSPHTEAAKDTGQNGP--KPPATLGADGPPP 780
                                                                                                                                                       869
                                                                                                                                                                                                                                                                                              662
                                                                                                                                                                                                                                                                                                                                                         870 AEEGPAKGKDAEAAEATA---EGALKAEKKEG------GSGRATTAKSSGAPQDSDSS 918
                                                                                                                                                                                                                                                                                                                                                                                           663 EEKPQEKEKEKEKVEEVGGKEEGGLKESRKEDIAINGEVEGKEEEQETKEKGSGGEEEKG 722
                                                       486 PKEELAAEAKVEKPEKAKSPVAKSPTTKSPTAKSPEAKSPEAKSPTAKSPT---AKSPVA
                                                                                                                                                                                                                                    BKEEETAAAPPVEEGE--EQKPPAAEE--LAVDTGKAEEPVKSEC-----TEE
                                                                                                                                                                                                                                                                                           603 BEKKEAKESPKEEKAEKKEEKPKDVPEKKKAESPVKAESPVKEEVPAKPVKVSPEKEAKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A. White Machael SER7267; MEDIATE=90067946; PubMed=2587267; MACHAED M.; Lee B., Ramirez F., Machado M., Horton W.; Nucleotide sequence of the full length cDNA encoding for human type
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Vuorio B.; "Construction and identification of a cDNA clone for human type II
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        pro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 collagen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Euteleostomi;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      F.G., Solomon
type II collag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE OF 1106-1418 FROM N.A.
MEDLINE=8806771; PubMed=2825137;
Elima K., Vuorio T., Vuorio B.;
"Determination of the single polyadenylation site of the human alpha 1(II) collagen gene.";
Nucleic Acids Res. 15:9499-9504(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        21-JUL-1986 (Rel. 01, Created)
01-JAN-1990 (Rel. 13, Last sequence update)
15-WAR-2004 (Rel. 43, Last annotation update)
Collagen alpha 1(II) chain precursor [Contains: Chondrocalcin]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               gene.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Elima K., Maekelae J.K., Vuorio T., Kauppinen S., Knowles
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Craniata; Vertebrata; Eutele
Catarrhini; Hominidae; Homo
                                                                                                                  781 GPPTPPRRTSRAPIEPTPASEATGAPTPPPAPPSPSAP-PPVVPK-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Submitted (DEC-1988) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE=87031574; PubMed=3021582; Nunez A.M., Kohno K., Martin G.R., Yamada Y.; Nunez A.m., Fronto K., Martin G.R., Yamada Y.; Promoter region of the human pro-alpha l(II)-collagen Gene 44:11-16(1986).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE OF 963-1418 FROM N.A. MEDLINE-85190534; PubMed-3857598; Cheah K.S. E., Stoker N.G., Griffin J.R., Grosveld "Identification and characterization of the human
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Proc. Natl. Acad. Sci. U.S.A. 82:2555-2559(1985)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT; 1418 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   II procollagen.";
Nucleic Acids Res. 17:9473-9473(1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        919 ATCSADEV---DEAEGGDKN 935
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE OF 1120-1398 FROM N.A. MEDLINE-85306861; PubMed=3840017;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Biochem. J. 229:183-188(1985).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Eukaryota; Metazoa; Chordata;
Mammalia; Eutheria; Primates;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE OF 432-1145 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE OF 1-28 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     procollagen mRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (COL2A1)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HUMAN
732
                                                                                                                                                                                                                                    826
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ramirez
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            P02458;
```

į

```
MEDLINE=86104139; PubMed=3002437;
Nunez A.M., Francomano C., Young M.F., Martin G.R., Yamada Y.;
"Isolation and partial characterization of genomic clones coding for a human pro-alpha 1 (II) collagen chain and demonstration of restriction fragment length polymorphism at the 3' end of the gene.";
Blochemistry 24:6343-6348(1985).
                                                                                                                                                                                                                                                                                               SEQUENCE OF 35-167 FROM N.A.
MEDLINE-89233138; PubMed=2714801;
Su M.W., Benson-Chanda V., Vissing H., Ramirez F.;
"Organization of the exons coding for pro alpha 1(II) collagen N-
propeptide confirms a distinct evolutionary history of this domain of
the fibrillar collagen genes.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              rare and some common diseases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Kuivaniemi H., Tromp G., Prockop D.J., "Mutations in fibrillar collagens (types I, II, III, and XI), fibrilassociated collagen (type IX), and network-forming collagen (type X) cause a spectrum of diseases of bone, cartilage, and blood vessels.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Bateman J.F., Hannagan M., Chan D., Cole W.G.;
"Characterization of a type I collagen alpha 2(I) glycine-586 to
valine substitution in osteogenesis imperfecta type IV. Detection of
the mutation and prenatal diagnosis by a chemical cleavage method.";
Biochem. J. 276:765-770(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     "Glycine to serine substitution in the triple helical domain of praalpha 1 (II) collagen results in a lethal perinatal form of short-limbed dwarfism.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (COL2A1) as
                                                                                                                                                                                                     Strom C.M., Uphoit W.B.; "Isolation and characterization of genomic clones corresponding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VARIANT OSTEOARTHRITIS CYS-650.
MEDLINE=90370826; PubMed=1975693;
Ala-Kokko L., Baldwin C.T., Moskowitz R.W., Prockop D.J.,
"Single base mutation in the type II procollagen gene (COI
cause of primary osteoarthritis associated with a mild
chondrodysplasia.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Vissing H., D'Alessio M., Lee B., Ramirez F., Godfrey M., Hollister D.W.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             VARIANT SEDC 1095-GLY--TYR-1330 DEL.
MEDLINE=89266907; PubMed=2543071;
Lee B., Vissing H., Ramirez F., Rogers D., Rimoin D.;
"Identification of the molecular defect in a family with spondyloepiphyseal dysplania.";
Science 244:978-980(1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Proc. Natl. Acad. Sci. U.S.A. 87:6565-6568(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Kuivaniemi H., Tromp G., Prockop D.J., "Mutations in collagen genes: causes of
                                                                                                                                                                                                                                              the human type II procollagen gene."; Nucleic Acids Res. 12:1025-1038(1984)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Biol. Chem. 264:18265-18267(1989).
                                                                                                                                                                                       MEDLINE=84118798; PubMed=6320112;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE=91184577; PubMed=2010058;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE=97255959; Pubmed=9101290;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE=90036909; PubMed=2572591;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VARIANT O1-IV VAL-717.
MEDLINE=91291136; PubMed=2064612;
                     SEQUENCE OF 1227-1289 FROM N.A.
                                                                                                                                                                    SEQUENCE OF 1176-1226 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mutat. 9:300-315(1997)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FASEB J. 5:2052-2060(1991).
                                                                                                                                                                                                                                                                                                                                                                                                              Genomics 4:438-441 (1989)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               [11]
REVIEW ON VARIANTS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                   REVIEW ON VARIANTS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 VARIANT SER-1074.
```

pro-

Ø

VARIANT OSTEOARTHRITIS CYS-650. MEDLINE=91086471; PubMed=1985108;

```
MEDLINE=94063862; PubMed=8244341; Milliams C.J., Considine E.L., Knowlton R.G., Reginato A., Neumann G., Harrison D., Buxton P., Jimenez S.A., Prockop D.J.; Spondyloepiphyseal dysplasia and precolous osteoarthritis in a family with an Arg75-->Cys mutation in the procollagen type II gene
Eyre D.R., Weis M.A., Moskowitz R.W.;
"Cartiage expression of a type II collagen mutation in an inherited form of osteoarthritis associated with a mild chondrodysplasia.";
J. Clin. Invest. 87:357-361(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE=93304428; PubMed=8317498; Koerkkoe J., Ritvaniemi P., Haataja L., Kaeaeriaeinen H., Kiviriko K.I., Prockop D.J., Ala-Kokko L.; Mutation in type II procollagen (COLZA1) that substitutes aspartate for glycine alpha 1-67 and that causes cataracts and retinal detachment: evidence for molecular heterogeneity in the Wagner Syndrome and the Stickler syndrome (arthro-ophthalmopathy).";
                                                                                                                                                                                                                                                   VARIANT HYPOCHONDROGENESIS SER-705.
MEDILINE-9225484; PubMed=1374906;
MEDILINE-9225484; PubMed=1374906;
Ramirez F., Vitale E., Lee B.;
"Characterization of a type II collagen gene (COL2A1) mutation dientified in cultured chondrocytes from human hypochondrogenesis.";
Proc. Natl. Acad. Sci. U.S.A. 89:4583-4587(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Holderbaum D., Malemud C.C., Moskowitz R.W., Haggi T.W., "Human cartilage from late stage familial osteoarthritis transcribes Lype II collagen mRNA encoding a cysteine in position 519."; Biochem. Biophys. Res. Commun. 192:1169-1174(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ΙΙ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  "A dominant mutation in the type II collagen gene (COL2A1) produces spondyloepimetaphyseal dysplasia (SEMD), Strudwick type."; Am. J. Hum. Genet. 53:A209-A209(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         "A mutation in the amino-terminal end of the triple helix of type collagen causing severe osteochondrodysplasia.";
Genomics 16:282-285(1993).
                                                                                          VARIANT HYPOCHONDROGENESIS GLU-984.
MEDLINE=33054548; PubMed=1429602;
Bogaert R., Tiller G.E., Wies M.A., Gruber H.E., Rimoin D.L.,
Cohn D.H., Byre D.R.,
"An amino acid substitution (Gly853-->Glu) in the collagen alpha
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Vikkula M., Ritvaniemi P., Vuorio A.F., Kaitila I., Ala-Kokko L.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDINE-93315508; PubMed-8325895; Chan D., Taylor T.K.F., Cole W.G.; "Characterization of an arginine 789 to cysteine substitution in alpha 1 (II) collagen chains of a patient with spondyloepiphyseal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             G.E., Weis M.A., Lachman R.S., Cohn D.H., Rimoin D.L.,
                                                                                                                                                                                                1(II) chain produces hypochondrogenesis.";
J. Biol. Chem. 267:22522-22526 (1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           J. Biol. Chem. 268:15238-15245(1993)
[25]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    VARIANT SEDC SER-1128.
MEDLINE=93140139; PubMed=8423604;
Cole W.G., Hall R.K., Rogers J.G.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VARIANT OSTEOARTHRITIS CYS-650.
MEDLINE=93282819; PubMed=8507190;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VARIANT SEMD ARG-285.
MEDLINE-93252400; PubMed-8486375;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Genet. 92:499-505(1993).
                                                                                                                                                                                                                                                                                                                                                                                                               VARIANT WS-II ASP-198.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VARIANT SEMD CYS-840.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VARIANT SEDC CYS-206.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              VARIANT SEDC CYS-920.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Peltonen L.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Eyre D.R.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               dysplasia.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Tiller
```

```
86;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ---PRGKSRSPAPPADKEAFAAEAQKLPGDPPCWTSGLPFPVPPREVIKASPHAPDPSAF 1075
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1251
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1302 HETAA------PKRTYDMMEGRVGRA----ISSASIEGLMGR-----AIPPER 1339
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1340 HSPHHLKEQHHIRG-SITQGIPRSYVEAQEDYLRREAKLLKREGTPPP-----PPSRDL 1393
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1394 T-EAYKTQALGPL-----KLKPAHEGLVATVKEAGRSIHEIPREELRHTPELPLAPR 1444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PLKEGSITQGTPLKYDTGASTTGSKKHD-----VRSLIGSPGR-----TFPPVHPLDVM 1493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ADARALERACYEESLKSRPGTASSSGGSIARGAPVIVPELGKPRQSPLTYEDHGAPFAGH 1553
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PYSEHAKAPVGPV-TMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEASV 1194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PSRL-----DRGR--EDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPP 1301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               521
                                                                                                                                                                                                                                   794
                                                                                                                                                                                                                                                                                                     EPTPASEATGAPTP-----PPAPPSPSAPPVVPKEEKEEETAAAPPVEEGEEQKPPAA 848
                                                                                                                                                                                                                                                                                                                                                                           EBLAVDTGKAEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKS 908
                                                                                                                                                                                                                                                                                                                                                                                                                                             SGAPODSDSSATCSADEVDEAEGGDKNRLLLSPRPSLLTPTGDPRANASPOKPLDLKQLKQ 968
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             671
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ------GVKGE-----SGSPGENGSP-GPMGPRGLPG 261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          77
                                                                                                                                                                                   421 KRG-ARGE-PGGVGPIGPPGERGAPGNRGFPGQDGLAGPKGAP------GERG
"The clinical features of spondyloepiphyseal dysplasia congenita resulting from the substitution of glycine 997 by serine in the alpha 1.111] chain of type II collagen."; J. Med. Genet. 30:27-35(1933).
                                                                                                                                                            721 AEALHASGNEV----DRGECSGPATVNNSSDTESIPSPHTEAAK-----DTGQNG-P
                                                                                                                                                                                                                                   -----RRTSRAPI
                                                                                                                                                                                                                                                                  GPRGRDGEPGTLGNPGPPGPPGPPGLGGNPAAQMAGGFDEKAGGAQLGVMQGPMGPM
                                                                                                                                                                                                                                                                                                                                        GPRGPPGPAGAPGPQGFQGNPGEPGEPGVSGPMGPR------GPP---GPPGKPG--
                                                                                                                                                                                                                                                                                                                                                                                                 ERGRIGPAGAAGAR----GNDGQPGPAGPPGPVGPAGGPGFPGAPGAKGEAGPTGARGPEG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            319 AQGPRGEPGTPGSPG-----PAGASGNPG-----TDGIP-----GAKGSAGAPGIA--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SYAPPGHPLPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ----ATGPLGPKGQTGKP-----GIAGFKGEQ-GPKGEPGPAGPGGAPGFAGEEG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LRGTALGSVPGGSITKGIPSTR-VPSDSAITYRGSIT--HGTPADVLYKGTITRIIGEDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PSGLAGPKGANGDPGRPGEPGLPGARGL-TGRPGDA-GPQGKVGPSGAPGEDGR--PGPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 522 GPQGARGOPGVMGFPGPKGA----NGEPGKAGEKGLPGAPGLRGLPGKDGETGAEGPPGP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ------GPPGPPGEGGKPGDQGV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ----QGPRGLPGTPGTDGPKGASGPAGPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GAQGPPGLOGMP--GERGAAGIAGPKGDRGDVGEKGPEGAPGKDGGRGLTGPIGPPGPAG
                                                                                                                           727; Gaps
                                                                                         Length 1418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     969 RAAAIPPIQVTKVHEPPREDAAPTKPAPPAPPPPQNLQPESDAPQQPGSS-
                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                --GAPGPPGPRG-------------------
                                                                                        Query Match 2.5%; Score 330.5; DB 1; Best Local Similarity 20.4%; Pred. No. 0.00052; Matches 378; Conservative 117; Mismatches 631;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PGEAGAPGLVGPRGERGFPGERGSPGAQGL
                                                                                                                                                                                                                                 KP-----PATLGADGP--PPGPPTPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         578 AGPAGERGEQGAPGPSGFQGLP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1136
                                                                                                                                                                                              18
                                                                                                                                                                                                                                 168
                                                                                                                                                                                                                                                                  78
                                                                                                                                                                                                                                                                                                     795
                                                                                                                                                                                                                                                                                                                                                                           849
                                                                                                                                                                                                                                                                                                                                                                                                              184
                                                                                                                                                                                                                                                                                                                                                                                                                                                 606
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1019
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1076
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               466
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             617
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1445
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   672
 REFER
                                                                                                                                                                                             a
                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                              ઠે
                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                            δ
                                                                                                                                                                                                                                ઠે
                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           상 원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

```
1928 VYPTLMEPVLLPKEAPRVARPERPRADTGHA-------FLAKPPARSGL 1969
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2030 PFSIQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDKGLPKHLEELDKSHLEGELRPK 2089
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1198 ESIRSPEGSKKNPAR------TCRDLKLCHPEWKSGDYMIDPNQGCTLDAM 1242
                                                                       LLRGVSGVDLYRSHIPLAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIRGY 1673
                                                                                                                                                                                                                         -----SQVPHL--PVLVPPTPGTPATAMDRL 1758
                                                                                                                                                                                                                                                                                                                                                                          1819 TTVEHAPIWRPGTEQSSGSSGSSGGGGGSSSRPASHSHAHQHSPISPRTQDALQQRPSVL 1878
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ---GVKG-------DRGETGAVGAPGAPGPPGSPGPAGPTGKQGDRGEAGAQGP 1027
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MGPS-----GPAGARGIQGPQGPRGDKGEAGEPGERGLKGHRGFTGLQGLPGPPGPSGD 1081
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1970 EPASSPSKGSEPRPLVPPVSGHATIARTPAKNLAPHHASPDPPAPASASDPHREKTQSK 2029
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1138 PPGPPGPPGPGIDMSAFAGLGPREKGPDPLQYMRADQAAGGLRQHDAEVDATLKSLNNQI 1197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1243 KVFCNMETGETCVYPNPANVPKKNWWSSKSKEKKHIWFGETINGGFHFSYGDDNLAPNTA 1302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ---ITYHCKNSIAYLDEAA 1335
LPRGSPVTMREPTPRLQEGSLSSSKASQDRKLTSTPREIAKSPHSTVPEHHPHPISPYEH 1613
                                                                                                                                                                                                                                                                                                AYLPTAPQPFSSRHSSSPLSPGGPTHLTKPTTTSSSERERDRDRERDRDREREKSILTST
                                                                                                                                                                                                                                                                                                                                                                                                                                                  HNTGMKGIITAVEPSKPTVLRSTSTSSPVRPAATPPPATHCPLG-----GTLDG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ----GPPGPVGPS-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2090 OPGPVKLGGEAAHLPHLRPLPESQPSSSPLL----QTAPGVKGHQRVV----TLAQHI
                                                                                                          --TGVTGPKGARG----PQGP---PGATGFPGAAGRVGPPGSNGNPG----PP----GP
                                                                                                                                                                                    -----GARGD --SGPPGRAGEPGLOGPAGPPGEKG
                                                                                                                                                                                                                                                            EPGDDGPSGAEGPPGPQGLAGQRGIVGLPGQRGERGFPGLPGPSGEPGQQGAPGASGDR-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2140 SEVITQDYTRHHPQQLSAPLPAPLYSFPGASCPVLDLRRP--PSDLYLPPPDHGAPARGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    P---HSEGGKRSPEPNKTSV-----------GEDGIEPVSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2228 PEGMTEPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQPPAFFSKLTESNSAMVKSKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2288 QEINKKINTHNRNEPEYNISQPGTEIFNMPAIT------GTGLMTYRSQ 2330
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GNLKKALLIQGSNDVE--IRAEGNSRFTYTALKDGCTKHTGKWGKTVIEYRSQ 1386
                                  --- AGAPGPQGPSGAPGPQGP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Homo sapiens (Human).
Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia, Eutheria, Primates, Catarrhini, Hominidae, Homo.
                                                                                                                                               PDTAALENROTIINDYITSQQMHHNTATAMAQRADMLRGLSPRESSLALNYAAGP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----GKDGANGIPGPIGPGPRGRSGETGPAGPPGN----
                                                                                                                                                                                                                                                                                                                                 sequence update)
annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1303 NVQMT-----FLRLLSTEGSQN-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT; 2349 AA
                                                                                                                                                                                                                                                                                                                                                                                                               ----GPAGEPGREGSPGADGPPGRDGAA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QGASGPAGPSGPR-------
                                  ---GOPGAKGEOGEAGÓKGD--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TPR_HUMAN STANDARD
P12270;
01-OCT-1989 (Rel. 12, C
01-OCT-1996 (Rel. 34, L
15-MAR-2004 (Rel. 43, L
Nucleoprotein TPR.
1554
                                                                       1614
                                                                                                                                                                                                                                                          889
                                                                                                                                                                                                                                                                                                                                                                                                                                                  1879
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1028
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1082
                                                                                                                                               1674
                                                                                                                                                                                                                         1729
                                                                                                                                                                                                                                                                                                1759
                                                                                                                                                                                    851
                                                                                                                                                                                                                                                                                                                                                                                                               960
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       984
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1336
                                774
                                                                                                          809
                                                                                                                                                                                                                                                                                                                                   948
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 77
TPR HUMAN
TD TPR HUMAN
AC 01-0C
DT 01-0C
DT 15-MP
DB Nucle
GN TPR.
GN TPR.
GS BUMON
OC BUMON
                                                                                                  g
                                                                                                                                           ò
                                                                                                                                                                                g
                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                셤
                                                                     ò
                                                                                                                                                                                                                         ઠ
                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
```

0.00083;

17.8%; Pred. No.

```
Best Local Similarity
Matches 305; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               517
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    627
                                                                                                                                                                                                                                                                                       897
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              461
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1367
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       658
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         681
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              733
                                                                                                           QQ
                                                                                                                                                                                                                                                                                  셤
                                                                         à
                                                                                                                                                       ò
                                                                                                                                                                                                d
                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                               ઠે
                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      or RAF genes.
DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
WWW="http://www.infobiogen.fr/services/chromcancer/Genes/TPRID282.html".
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Oncogene 2:617-619(1988).

-1- FUNCTION: Component of the cytoplasmic fibrils of the nuclear pore complex implicated in nuclear protein import. Its amino terminus is involved in activation of oncogenic kinases.

-1- SUBCELLULAR LOCATION: CYTOPLASMIC SURFACE OF THE NUCLEAR PORE COMPLEX. THE ASSEMBLY OF THE NPC IS A STEPWISE PROCESS IN WHICH TRP-CONTAINING PERIPHERAL STRUCTURES ASSEMBLE AFTER OTHER
                                                                                                                                                                                                             MEDLINE=95096166; PubMed=7798308;
Byrd D.A., Sweet D.J., Pante N., Konstantinov K.N., Guan T.,
Saphire A.C.S., Mitchell P.J., Cooper C.S., Aebi U., Gerace L.;
"Tpr, a large coiled coil protein whose amino terminus is involved in activation of oncogenic kinases, is localized to the cytoplasmic surface of the nuclear pore complex.";
J. Cell Biol. 127:1515-1526(1994)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMPONENTS, INCLUDING P62.

TISSUE SPECIFICITY: Highest in testis, lung, thymus, spleen and brain, lower levels in heart, liver and kidney.

DISEASE: Involved in tumorigenic rearrangements with the MET, TR
                                                                         Mitchell P.J.; Cooper C.S.;
"The human tpr gene encodes a protein of 2094 amino acids that has extensive coiled-coil regions and an acidic C-terminal domain.";
Oncogene 7:2329-2333(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AFDD6885CEDCA9EF CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                             H.W.S., Tempest P.R., Merrifield K.R., Rance A.J.; homologues activate met and raf.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      translocation;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL). (POTENTIAL). (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MIN; LOSYSUN; C:CYtOplasm; TAS.
GO; GO:0005737; C:CYtOplasm; TAS.
GO; GO:0005643; C:nuclear pore; TAS.
GO; GO:006666; P:protein-nucleus import; TAS.
Colled coll; Prote-oncogene; Chromosomal trans.
Nuclear protein; Transport.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COLLED COLL
COLLED COLL
COLLED COLL
COLLED COLL
COLLED COLL
COLLED COLL
COLLED COLL
COLLED COLL
COLLED COLL
POLY-SER.
POLY-SER.
POLY-SER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COIL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COILED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COILED
                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE OF 1-142 FROM N.A.
MEDLINE-88262257; PubMed=3387099;
                                                                MEDLINE=93064711; PubMed=1437155;
                                                                                                                                                                                           REVISIONS, AND CHARACTERIZATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; X66397; CAA47021.1; -.
EMBL; Y00672; CAA68681.1; -.
Genew; HGNC:12017; TPR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    628
805
869
979
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1304
1595
1595
530
1964
2298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AA;
                                            FROM N.A.
  NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MIM; 189940; -
                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                               King
"tpr
SO THE THE THE THE TENT TO BE SEED OF THE THE THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT TO THE TENT
```

Length 2349;

DB 1;

Score 330.5;

2.5%;

Query Match

```
327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         : |:: : | :| :| 1074
| KQELQDDKRRAIESMEQQLSELKKTLSSVQNEVQEALQRASTALSNEQQARRDCQEQAKI 1074
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1075 AVEAQNKYERELMLHAADVEALQAAKEQVSKMASVRQHLEETTQKAESQLLECKASWEER 1134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1250 QVTAKTWAQ---HEELMKKTETMNVVMETNKMLREEKERLEQDLQQMQAKVRKLELDILP 1306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RIQQLTEEIGRLKAEIARSNASLTNNQNLI------QSLKEDLNKVRTEKETIQKD 1416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             : ::| ::|
1592 KDELDVRITALKSQYEGRISRLERELREHQERHLEQRDEPQEPSNKVPEQQRQITLKTTP 1651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1652 ASGERGIASTSDPPTANIKPTPVVSTPSKVTAAAMAGNKSTPRASIRPMVTPATV---- 1706
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    .707 -TNPTTTPTATVMPTTQVESQEAMQSEGPVEHVPVFGSTSGSVRSTSPNVQPSISQPILT 1765
                                                                                                                                847
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    L-----DAKIIDIQEKVKTITQVKKIGRRYKTQYEELKAQQDKVMETSAQSSGDHQEQHV 1471
                                                                                                                                                                                                                                                                                                               177 DRVDREITMVEQQISKLKKKQQQLEEEAAKPPEPEKPVSPPFIESKHRSLVQIIYD--EN 234
                                                                                                                                                                                                                                                                                                                                                           --AQKEIATLKQHLSNMEVQVASQSSQRTGKGQPSNKEDVDDLVSQLRQTEEQVNDLKER 954
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             --KVREYYEKQF------PEIRKQRELQERMQSRVGQRGSGLSMSAARSEHE 371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1532 LODRTTQEEQLRQQITEKEEKTRKAIVAAKSKIAHLAGVKDQLTKENEELKQRNGALDQQ 1591
                                                                                                                                                                                                                                                   848 ISHLKKKL-----ENEVEORHTLTRNLDVOLLDTKRQLDTETNL---HLNTKELLKN- 896
                                                                                                                                                                                                                                                                                                                                                                                                                                                235 RKKAEAAHRILEGLGPQVELPLYNQPSDTRQYHENIKIN------QAMRKKLILYFKRR 287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QEBKDEKEKEKEAEKEEEKPEVEND-----KEDLLKEKTDDTSGEDN---DEKEAVAS 566
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        867
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ----SRAPIEPTPASEATGAPTPPPAP-PSPSAPPPVVP 823
                                                                                                                         SEVRISQORESLIAE-----ORGÓNI-LITNLOTIQGILERSETETKORLSSOIEKLEHE
                                                                                                                                                                                      117 LLRPSPLLATGQPAGSEDLTKDRSLTGKLEPVSPPSPPHTDPELELVPPRLSKEELIQNM
                                                         57 SIIQPQRRRPSLLSEFQPGNERSQELHLRPESHSYLPELGKSEMEFIESKRPRLELLPDP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ---NPRRRAKES---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ----KTVAECVLYYYLTKKNENYKSLVRRSYRRRGKSQQQQQQQQQQQQQQQQPMPRSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          567 KGRKTANSQGRRKGRITRSMANEANSEEAITPQQSAELASMELNESSRWTEEEMETAKKG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      372 VSEIIDGLSE----QENLEKQMRQLAVIPPMLYDADQQRIKFINMNGLMADPMKV----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ----RG--ECSGPATVN-NSSDTESIPSPHTEAAKDTGQNGPK-----PPATLGADGP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----EEMEASG----
Mismatches 612; Indels 569; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -YKDRQVMNM--WSEQEKETFREKF-----MQHPKNFGLIASFLER-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -----MEKERNAR------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 824 KEEKEETAAAPPVEEGEEOKPPAAEELA-----VDTGKAEEPVKSECT---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ---VSGNEEEMVEEAEALHASGNEVP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LLEHGRNWSAIARMVGSKTVSQCKNFYFNYK--------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             288 NHARKOWKOKFCORYDOLMEALEKKVERIEN-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative 226;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ---KRQNLDEILQQHKLK---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        779 PPGPPTPPRRT-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       =
::
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
```

```
320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     842
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                     "Gene
셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   장;, 음
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2206
                                                                                         1885
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2100
1766 VQQQTQATAFVQPTQQSHPQ1EPANQELSSNIVEVVQSSPVERPSTSTAVFGTVSATPSS 1825
                                                                                                                                                                                                                                                                                                                                                                                                                   QNLQPESDAPQQPGSSPRGKSRSPAPPADKEAFAAEAQKLPGDPPCWTSGLPFPVPPREV 1062
                                                                                                                                                                                                                                                                                                                                                                                                                                                              ON-SGEGNTGAAESSFSQEVSREQOPSSASEROAPRAPOSPRRP-----PHPLPPRLT 2072
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1176 GQAGPPESLGVPTAQEASVLRGTALGSV------PGGSITKGIPSTRVPSDS 1221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AITYRGSITHGTPADVLYKGT----ITRIIGEDSPSRLDRGREDSLPKGHVIYEGKKGH 1276
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1277 VLSYEGGMSVTQCSKED-----GRSSSGPPHETAAPKRTYDMMEGRVGRAISSAS 1326
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2260 GISSEAGLEIDSQQEEEPVQASDESDLPSTSQDPPSSSSVDTSSSQPKPFRRVR-LQTTL 2318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              943 SLLTPTGDPRANASPQKPLDLKQLKQRAAAIPPIQVTKVHEPPREDAAPTKPAPPAPPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1826 SLPKRTREEEEDSTIEASDQVSDDTVEMPLPKKLKSVTPVGTEEEVMAEESTDGEVETQV
                                                                                                                                                                                1886 YNQDSQDSIGEGVTQGDYTPMEDSEETSQSLQIDLGPLQSDQQTTTSSQDGQGKGDDVIV
                                                                                                                                                                                                                                -----SSATCSAD----EVDEAEGGDKNRLLSPRP
                                                                                                                                                                                                                                                                         1946 IDSDDEEEDEEDDDDDEDDTGMGDEGEDSNEGTGSADGNDGYEADDAEGGDGT----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IKASPHAPDPSAFSYAPPGHPLPLGLHDTARPVLPRPPTISNPPPLISSAKHPSV----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ------POELGPPVORIOMTRROSVGRGLO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LERQIGAISQCMSVQLHVPYSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLS--PR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LTPGIGGMQQ-----HFFDDEDRTVPSTP-TLVVPHRTDGFAEA----IHSPQVAGVPR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE=97313266; PubMed=9168970; Churcher C.M., Bowman S., Badcock K., Bankier A., Brown D., Churcher C.M., Bowman S., Badcock K., Bankier A., Brown D., Chilingworth T., Connor R., Devlin K., Gentles S., Hamlin N., Harris D.E., Horsnell T., Hunt S., Jagels K., Jones M., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A., Rice P., Rowley N., Skelton J., Smith V., Walsh S., Whitehead S., Barrell B.G., "The nucleotide sequence of Saccharomyces cerevisiae chromosome IX."; Nature 387:84-87(1997).
                                                                                                                                    ---RATTAKSSGAPO---
                                             --ATAEGALKAEKKEG--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Eukaryota; Fungi, Ascomycota, Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AWYH YEAST STANDARD; PRT; 1367 AA.
P08640; P08068;
01-AUG-1988 (Rel. 08, Created)
01-FBB-1995 (Rel. 31, Last sequence update)
10-OCT-2003 (Rel. 42, Last annotation update)
Glucoamylase S1/S2 precursor (RC 3.2.1.3) (Glucan 1,4-alpha-glucosidase) (1,4-alpha-D-glucan glucohydrolase).
STA2 OR MALS OR YIR019C.
                                                                                                                                                                                                                                                                                                                                                                     -----DP--GTETEESMGGGEGNHRAA------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1327 IEGLMGRAIPPERHSPHHLKEQHHIRGSITQG 1358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2319 ROGVRGROFINRORGVSHAMGG----RGGINRG 2346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Saccharomyces cerevisiae (Baker's yeast).
                                           ----EEAEEGPAKGKDAEAAE----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
STRAIN=S288c / AB972;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            I----HAP----
                                                                                                                                                                                                                              -DSD----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=4932;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2101
                                                                                                                                    899
                                                                                                                                                                                                                                                                                                                                                                     1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                2021
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1063
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2073
                                             868
                                                                                                                                                                                                                                                                                                                                                                                                                   1003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1222
                                                                                                                                                                                                                              914
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 78
AMTH YEAST
1D 74NTH YEAST
DT 01-4NG
DT 01-7NG
DE Glucoa
DE Glucoa
DE Glucoa
DE Glucoa
OC Saccha
OC Saccha
OX NCBI T
RN III
RP SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
RY SEQUENT
                                                                                    셤
                                                                                                                                                                      В
                                                                                                                                                                                                                                                                                                                                                           유
                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            쉽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          용
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 · 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                         δ
                                                                                                                                  ð
                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                        ठे
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ठ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              58;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RANASPQKPLDLKQLKQRAAAIPPIQVTKVHEPPREDAAPTKPAPPPPPPQNLQPESDA 1011
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 784
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                741
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      841
                                                                                                                                                                                                                                                                                                                                                                                               from Saccharomyces cerevisiae.";
FBES Lett. 239.179-184(1988).
-!- CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-glucose residues successively from non-reducing ends of the chains with release of beta-D-glucose.
-!- SIMILARITY: TO S.POMBE SPEC15.13.
-!- SIMILARITY: SOME, TO S.POMBE SPCC285.13C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               209 KSSTTTSSTSESSTTTSSTSESSTTTSSTSESSTTTSSTSESSTSSST-----TAPAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STTESSSAPV-PTPSSSATTESSSAPVTSSTTESSSAPVPT--PSSSTTESSSAPVTSSTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STTESSSAPVTSSTTESSSAPVPTPSSST-----TESSSAPVTSSTTESSSAPVPTPSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 742 VNNSSDTESIPSPHTEAAKDTGQNGPKPP--------ATLGADGPPPGPPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                262 ртттscткекртрртттs--сткекртррнногтрстккиттскисткиттруртрss
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      785 PPRRTSRAPIEPTPAS---EATGAPTPPPAPPSPSAPPPVVPKEEKEEETAAAPPVEEGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EQKPPAAEELAVDTGKAEEPVKSECTEEAEEGPAKGKDAEAAEA-----TAEGALKA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            894 EKKEGGSGRAT--TAKSSGAPODSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                682 KKKKAPAAASEEAAFPPVVEDEEMEASGVSGNEEEMVEEAEALHASGNEVPRGECSGPAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ESSSAPVISSTIESSSAPVPIPSSSTIESSSAPVISSTIESSSAPVISSTIESSSAPVIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                 Pardo J.M., Ianez E., Zalacain M., Claros M.G., Jimenez A.; "Similar short elements in the 5' regions of the STA2 and SGA genes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GO; GO:0005886; C:plasma membrane; IDA.
GO; GO:0010447; P:filamentous growth; IDA.
GO; GO:0007125; P:invasive growth; IMP.
GO; GO:0007124; P:pseudohyphal growth; IMP.
Hydrolase; Glycosidase; Polysaccharide degradation; Glycoprotein;
                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED (GLCNAC. . .) (POTENTIAL)
N-LINKED (GLCNAC. . .) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 330; DB 1; Length 1367;
Pred. No. 0.00052;
1; Mismatches 601; Indels 380;
                                                                                                 underlying the evolution
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MW; 91C00E2DBD61AA9D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GLUCOAMYLASE S1/S2
SER/THR-RICH.
SEQUENCE OF 1-242 AND 762-1331 FROM N.A.
                                                                  nita I., Nakamura M., Fukui S.;
fusion is a possible mechanism
                                                                                                                                                                                                                                                                  STRAIN=SPX101-1C;
MEDLINE=89031230; Pubmed=3141213;
                                                                                                                                                                Bacteriol. 169:2142-2149(1987)
                               MEDLINE=87194600; PubMed=3106330;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 2.5%; SCO
Best Local Similarity 19.5%; Pre
Matches 277; Conservative 161;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL, Z38061; CAA86176.1; -.
EMBL, M16164; AAA35014.1; -.
EMBL, X13857; CAA320501; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                136110
                                                                                                                                                                                                                                  SEQUENCE OF 1-31 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Signal; Multigene family. SIGNAL 1 21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PIR; S48478; S48478. GermOnline; 139731; -. SGD; S0001458; MUC1. GO; GO:0005886; C:plasm
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          817
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          874 87
1367 AA;
                                                                  Yamashita
```

																			_													 .	<u> </u>	-	
491 STTESSSAPVTSSTTESSSAPVPTPSSSTTESSSA-PAPTPSSSTTESSSA 540	1012 PQQPGSSPRGKSRSPADPADKEAFAAEAQKLPGDPPCWISGLPFPVP 1058		1059 PREVIK-ASPHAPDPSAFSYAPPGHPLPLGLHDTARPVLPRPPTISNPPPLISSA 1112	600 SSSTIESSSAPAPIPSSSTIESSSAPVISSTIESSSAPVPTSSSTIESSSAPVPIPSS 658	1113 KHPSVLERQIGAISQCMSVQLHVPYŞEHAKAPVGPYTMGLPLPMDPKKLAPFSGVKQEQL 1172	659STIESSSAPVPIPSSSITIESSSAPVISSTIESSSAPVISSITESS 703	1173 SPRGQAGPPESLGVPTAQEASVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRG 1227	704 SAPVPIPSSSTIESSSAPVPIPSSSTIESSSAPVPIPSSSTIESSSAPVISSTIESS 760	1228 SITHGTPADVLYKGTITRIIGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMGVT 1287	761 SAPVPTPSSSTT 772		773 ESSSAPVPTPSSSTTESSSAPVPTPSSSTTESSVAPVPTPSSSS 816	1347 EQHHIRGSITQGIPRSYVEAQEDYLRREAKLLKREGTPPPPPFSRDLTEAYKTQALGP 1404	017SSAPSSTPFSSSTESSSVPVPTPSSSTTESSSAP 853	1405 LKLKPAHEGLVATVKEAGRSIHEIPREELRHTPELPLAPRPLKEGSITOGTPLKYDTGAS 1464	854 VS-SSTTESSVAPV	TTGSKKHDVRSLIGSPGRTFPPVHPLDVMADARALBRACYB	888 TTESFSTGTTVTPSSSKYPGSQTETSVS-ST 917	1525 GAPVIVPELGKPRQSPLTVEDHGAPFAGHLPRG-SPVTWREPTPRLQEGSLSS 1576	918 TETTIVPTKTTTSVTTPSTTTTTTVCSTGTNSAGETTSGCSPKTVTTTVPTTTTTSVTT 977	1577 SKASODRKLTSTPREIAKSPHSTVPEHHPHPISPYEHLLRGVSGVDLYRSHIPLAFDPTS 1636	978 SSTTIITTVCSTGTNSAGETTSGCSPKT 1006	1637 IPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAALBNRQTIINDYITSQQMH 1696		1697 HNTATAWAQRADMLRGLSPRESSLALNYAAGPRGIIDLSQVPHLPVLVPPTPGTPATAMD 1756	1054 GEITTTFVTKNIPTTYLTIAPTPSVTTVTNFTPTTTT 1092	1757 RLAYLPTAPQPFSSRHSSSPLSPGGPTHLTKPTTTSSSFRERDRDRERDRDREREKSILT 1816	1093 TVCSTGTNSAGBTTSGCSPKTVT-TTVPCSTGTGBYTTBATTLVT 1136	1817 STTTVEHAPIWRPGTEQSSGSSGSSGGGGGSSSRPASH-SHAHQHSPISPRYQDAL 1871		1872 QORPSVLHNTGMKGIITAVEPSKPTVLRSTSTSSPVRPAATFPPATHCPLGGTL 1925		1926 DGVYPTLMEPVLLPKEAPRVARPERPRADTGHAFLAKPPAR-SGLEPASSPSKGSEP-RP 1983	1240TTVITTES-SVGTNSAGETTTGYTTKSIPTTYITTLIPGSNGAKNYETVAT 1289	1984 LVPPVSGHATIART-PAKNIAPHASPDPPAPASAS 2019
QQ	δ	qq	ò	g	δ	QQ	ò	ор	ò	QQ Q	ò	qq	ò	Op Op	ò	Ωp	ò	අු	<u>ک</u>	ф	ò	QQ	δ,	qq	δλ	qq	δλ	qq	δ	QQ	ò	qq	'n	QQ	È

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EWBL outstation the EwBL outstation and the EwBL outstation on its use by non-profit institutions as long as its content is in oway modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ibb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           86;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           722
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   69 AAVLAAAAGATRPPRPPSAQQQRHARRG--SGEIVVLDDEDEEE-DEPGSPAAGSPGRAL 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 780 ---PGPPTPPRRISRA---PI-EPTPASEATGAPTPPPAPPSPSAPP----PVVPKEEKE 828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           68
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            673 MEKERNARRKKKKAPAAASEEA-AFPPVVEDEEMEASGVSGNE------EEMVEEAE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              723 A---LHASGNEVP-----RGECSGPATVNNSSDTESIPSPHTEAAKDTGQ--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 2.5%; Score 330; DB 1; Length 1461;
Best Local Similarity 21.0%; Pred. No. 0.00056;
Matches 366; Conservative 128; Mismatches 602; Indels 644; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       765 ------NGPKPP----ATLGADGPP-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Cheung A.K.;
Submitted (NOV-1989) to the EMBL/GenBank/DDBJ databases.
-!- FUNCTION: THIS IE PROTEIN IS A MULTIFUNCTIONAL PROTEIN CAPABLE
OF MIGRATING TO THE NUCLEUS, BINDING TO DNA, TRANS-ACTIVATING
OTHER VIRAL GENES, AND AUTOREGULATING ITS OWN SYNTHESIS.
-!- SUBCELLULAR LOCATION: Nucleus of infected cells.
-!- FUN: A long stretch of serine residues may be a major site of
phosphorylation.
-!- SIMILARITY: BELONGS TO THE ICP4/IE140/IE180 FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Cheung A.K.;
"DNA nucleotide sequence analysis of the immediate-early gene of pseudorabies virus.";
Nucleic Acids Res. 17:4637-4646 (1989).
                                                                                                                                                                                                                                                                                                                                                                                                          Pseudorabies virus (strain Indiana-Funkhauser / Becker) (PRV). Viruses; dSDNA viruses, no RNA stage; Herpesviridae; Alphaherpesvirinae; Varicellovirus. NCBI_TaxID=31523;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR005205; Herpes ICP4 C.
InterPro; IPR005206; Herpes ICP4 N.
Pfam; PF03585; Herpes ICP4 C; 1.
Pfam; PF03584; Herpes ICP4 N; 1.
Barly protein; Transcription regulation; Trans-acting factor;
DNA-binding; Phosphorylation; Nuclear protein.
POMAIN 390 405
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POLY-SER.
W; 7F31E7ABE403B208 CRC64;
1290 ATNPISIKTTSQLATTASASSVAPVVTSPSLTGPLQSAS 1328
                                                                                                                                                                                                                                                                              Last sequence update)
Last annotation update)
                                                                                                                                                                              PRT; 1461 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 <u>:</u>
=
::
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
MEDLINE=89315207; PubMed=2546124;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1461 AA; 149833 MW;
                                                                                                                                                                                                                                      01-OCT-1989 (Rel. 12, Created)
01-APR-1990 (Rel. 14, Last seq
01-FEB-1994 (Rel. 28, Last ann
                                                                                                                                                                                                                                                               01-APR-1990 (Rel. 14, Last seq
01-FEB-1994 (Rel. 28, Last ann
Immediate-early protein IE180.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; X15120; CAA33214.1; -.
                                                                                                                                                                              STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PIR; S04713; EDBEIF.
                                                                                                                                                                          PRVIF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REVISIONS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                          IE18 PR
P11675;
                                                                                                                                          IE18 PRVIF
                                                                                                                                                                                                      A PART TEN TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE 
   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            엄
```

8686868686	R S S C C C C C C C C C C C C C C C C C
186 SGSPGPSAAPRRWSPARGDPVGEPGPAARPR-TFAPPAQPAAVAAAPARRGPASP	1331VOLHVPYSEHAKAPUGPUTGLPLPMDPKGLAP-F8GYKQEQLSPRGQ-
8 6 8 6 8 6 8 6 8 6 8 6 8 6 8	8 6 8 6 8 6 8 6 8 6 8 6 8 6 8 6 8 6 8 6

1138 QGVLLLSARDLGFAGAVEYLCSRLGAARRR------LIVLDTIEDWPADGPA---- 1183 1835 SGSSGSSGGGGSSSRPASHSHAHQHSPISPRTQDAL-----QQRPSVLHNTGMKG--I 1886 | : : | | : : | : : | | 1.3 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | | 1.5 | 1970 EPASSPSK------GSEPRPLVPPVSGHATIARTPAK--NLAPHHASPDPPA-- 2013 1338 GPEGLPAELRAFCAAALLEPDAEAAPLVLAPGALAAAGAPPAVRWDFAPFETSVRAAAGG 1397 2014 ----PPASASDPHREKTQSKPFSIQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDK 2068 1737 VPHLPVLVPPTPGTPATAMD-RLAYLPTAPQPPSSR-----HSSSPLSPGGPT--HLTK 1787 1788 PTTTSSSERE------RDRDRERDRDREREKSILTSTTTVEHAPIWRPGTEQS 1834 neurofilaments.";
Proc. Natl. Acad. Sci. U.S.A. 85:1998-2002(1988).
-!- FUNCTION: Neurofilaments usually contain three intermediate
-!- FUNCTION: Neurofilaments usually contain three intermediate
filament proteins: L, M, and H which are involved in the
maintenance of neuronal caliber.
-!- FTM: There are a number of repeats of the tripeptide K-S-P, NFM is
phosphorylated on a number of the serines in this motif. It is
thought that phosphorylation of NFM results in the formation of
interfilament cross bridges that are important in the maintenance JEQUENCE FROM N.A.

(EDLINE-87275853; PubMed-3608989;

(Yyers M.W., Lazzarini R.A., Lee V.M.-Y., Schlaepfer W.W., Nelson D.L.;

(The human mid-size neurofilament subunit: a repeated protein

sequence and the relationship of its gene to the intermediate PTM: Phosphorylation seems to play a major role in the functioning of the larger neurofilament polypeptides (NF-M and NF-H), the levels of phosphorylation being altered developmentally and coincident with a change in the neurofilament function. SIMILARITY: Belongs to the intermediate filament family. 1887 ITAVEPS--------KPTVLRSTSTSSPVRPAATFPPATHCPLGGTLDGVYPT 1932 L--MEPVLLPKEAPRVARPERPRADTGHAFLAKPPARSGL---------[2] EQUENCE OF ONE OF THE 13 RESIDUR REPEATS.

REDLINE-88158120; PubMed=2450354;

REDLINE-88158120; PubMed=2450354;

REDLINE-88158120; PubMed=2450354;

Redline R. A.; Identification of the major multiphosphorylation site in mammalian Sukaryoča, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. (CBI_TaxID=9606; 01-APR-1988 (Rel. 07, Created)
01-AUG-1988 (Rel. 08, Last sequence update)
115-MAR-2004 (Rel. 43, Last annotation update)
Neurofilament triplet M protein (160 kDa neurofilament protein)
(Neurofilament medium polypeptide) (NP-M) (Neurofilament 3).
NEF3 OR NEFM OR NFM. 915 AA. PRT; ilament gene family."; MBO J. 6:1617-1626(1987). STANDARD; of axonal caliber. lomo sapiens (Human) NFM_HUMAN P07197;

```
(Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Homo sapiens
               373
                                                                                                                                               546
                                                                                                                                                                                        580
                                                                                                                                                                                                             999
                                                                                                                                                                                                                                                      725
                                                                                                                                                                                                                                                                            628
                                                                                                                                                                                                                                                                                                785
                                                                                                                                                                                                                                                                                                                     629
                                                                                                                                                                                                                                                                                                                                          840
                                                                                                                                                                                                                                                                                                                                                                                   871
                                                                                                                                                                                                                                                                                                                                                                                                                                                   825
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 81
T2D3_HUMAN
                                                     셤
                                                                                                                                            ద
                                ર્જ
                                                                              à
                                                                                               셤
                                                                                                                       à
                                                                                                                                                                  ò
                                                                                                                                                                                   g
                                                                                                                                                                                                           ò
                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                       염
                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                             40;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DPLLRPSPLLATGOPAGSEDLTKD-----RSLTGKLEPVSPPSPPHTDPELELVPPRLS 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          KEBLIQNMDRVDREITMVEQQISKLKKKQQQLEEERAKPPEPEKPVSPP-----IESKH 223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       :|:: ::| : | : | : | : | EBDIHRLKERFEBEARIRDDTEAAIRALRKDIEEASLVKVELDKKVQSLQDEVAFLRSNH 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PGSIIQPQRRRPSLLSEFQPGNERSQELHLRPESHSYLPELGKSEMEFIESKRPRLELLP 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RS----LVQIIYDENRKKAEAAHRILEGLGPQVELPLYNQPSDTRQYHENIKINQAM-- 276
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -RKKLILYFKRRNHARKQWKQKFCQRYDQLMEALEKKVERIENNPRRRAKESKVREYYEK 335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----QFPEIRKQRELQERMQSRVGQRGSGLSMSAARSEHEVSEIIDGLSEQEN---- 384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LOSKSIELESVRGTKESLERQLSDIEE-----RHNHDLSSYQDTIQQLENELRGT 372
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ---LEKQMR------QLAVIPPMLYDADQQRIKFINNNGLMADPMKVYKDRQ 427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BEEVADLLAQI------QASHITVE-------RKDYLKTDISTALKE 271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SPSSGFRSQSWSRGSPSTVSSSYKRSMLAPRLAYSSAMLSSAESSLDPSQSSSLLNGGSG 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SGSTQLVAQTWRATEPRYPPHSLSYPVQIARTHTDVGLLEYQHHSRDYASHLS-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                             ilarity 19.2%; Pred. No. 0.00037;
Conservative 149; Mismatches 361; Indels 333;
                                                                                                                                                                                                                                                                                                                                                                                         O-LINKED (GLCNAC) (BY SIMILARITY)
O-LINKED (GLCNAC) (BY SIMILARITY)
                                                                                                                                        GO; GO:0005883; C:neurofilament; TAS.
GO; GO:0005200; F:structural constituent of cytoskeleton; TAS
                                                                                                                                                                                                                            PROSITE; PS00226; IF; 1.
Intermediate filament; Coiled coil; Neurone; Phosphorylation;
                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 1; Length 915;
                                                                                                                                                                                                                                                                                                                                                                                                                 SFSD3DF34C9D9E50 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                  6 X 13 AA TANDEM REPEATS.
                                                                                                                                                                                                                                                                                                                                                                                                                                   2.5%; Score 329.5; DB 119.2%; Pred. No. 0.00037;
                                                                                                                                                                                                                                                                                                                          COIL 1B.
LINKER 12.
COIL 2A.
LINKER 2.
COIL 2B.
                                                                                                                                                                                                                                                                                                       COIL 1A.
LINKER 1.
                                                                                                                                                           InterPro; IPR006821; Filament head
InterPro; IPR001664; IF.
InterPro; IPR002957; Keratin I.
Pfam; PF00038; filament; 1.
Pfam; PF04732; filament head; 1.
PRINTS; PR01248; TYPE1KERATIN.
                                                                                                                                                                                                                                                                                                                                                                                                                 Σ
                                                                                             EMBL; Y00067; CAA68276.1; -. PIR; A27864; A27864. Genew; HGNC:7734; NEF3. MIM; 162250; -.
                                                                                                                                                                                                                                                                                                                                                                                                               102317
                                                                                                                                                                                                                                                                                 411
915
135
                                                                                                                                                                                                                                                                                                                                                                                         46
430
915 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                              Local Similarity
                                                                                                                                                                                                                                                  Glycoprotein.
INIT MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                        200;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              277
                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  385
                                                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Loca
Matches
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8
```

```
428 VMNWWSEQEKETFREKFMQHPKNFGLIASFLERKTVAE--CVLYYYLTKKNENYKSLVRR 485
                                                                                                                                               430 TISSKIQKTKVEAPKLKVQHK----FVEBIIEFTKVEDEKSEMEBALTAITELAASMKE 485
                                                                                                                                                                                                                                   544
                                                                                                                                                                                                                                                             604
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       664
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              784
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       658
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ILQQHKLKMEKERNARRKKKKAPAAASEEAAFPPVVEDEEMEASGVSGNEEEMVEEAEAL 724
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         824
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             605 ASMELNESSRWTEEEMETAKKGLLEHGRNWSAIARWYGSKTVSQCKNFYFNYKKRQNLDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    486 SYRRRGKSQQQQQQQQQQQQQQQQQQQQPMPRSSQEEKDEK-EKEKEAEKEEEKPEVENDKED
                                                                                                                                                                                                                                                                                                                                                 545 LLKEKTDDTSGEDNDEKEAVASKGRKTANSQGRRKGRITRSMANEANSEEAITPQQSAEL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HASGNEVPRGECSGPATVNNSSDTESIPSPHTEAAKDTGQNGPKPPATLGADGPPPGPPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----PVPK------SPVEEKGKSPVEEKGKSPVEEKGKSPVEEKGKSPVPKS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PPRRISRAPIEPTPASEATGAPTPPPAPPSPSAPPV---VPKEEKEETAAAPPVE--E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    719 KKĖEKPKDVPĖ----KKKAĖSPVKEĖAVAĖVVTITKSVKVHLEKETKEEGKPLQQEKEKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----EEGPAKG----KDAEAAEATAEG--ALKAEKKEGGSGRATTAKSSGAPQD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           775 KAGGEGGSEBEGSDKGAKGSRKEDIAVNGEVEGKEEVEQETKEKGSGR-----EE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AF-28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
MEDLINE=97336072; PubMed=9192867;
MEDLINE=97336072; PubMed=9192867;
Medgus G., May M., Carre L., Chambon P., Davidson I.;
"Human TAPF(II)135 potentiates transcriptional activation by the AF-
of the retinoic acid, vitamin D3, and thyroid hormone receptors in
mammalian cells.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
MEDLINE=21638749; PubMed=11780052;
MEDLINE=2.1638749; PubMed=11780052;
Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L., Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M., Beaseley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              T2D3 HUMAN STANDARD; PRT; 1083 AA. 000268; Q99721; Q98R42; C000268; Q99721; Q98R42; Created) 36. Created) 15-JUL-1998 (Rel. 36. Last sequence update) 15-JUL-1998 (Rel. 36. Last sequence update) 15-JUL-1998 (Rel. 37. Last sequence Trib 15-JUL-1998 (Rel. 37. Last sequence Trib 135 kDa subunit (TAFII-135) (TAFII-130) (TAFII-130) (TAFII-130) (TAFII-130) (TAFII-130)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Euteleostomi;
                                                                                                                                                                                                                                                                                                                                                                                                                          --EAEGEEA-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Butel.
Mammalia, Butheria, Primates, Catarrhini, Hominidae, Homo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----VEEKGKS
                                                                                                                                                                                                                                                                                                                                                                                            GEEOKPPAAEELAVDTGKAEEPVKSECTEEA-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               915 SDSSATCSAD -- EVDEAEGGDKN 935
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EKGVVTNGLDLSPADEKKGGDKS 847
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ---EAKEEKKVEEKSEEVATK----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    600 LVAD--AKVEKPEKAKSPVPKSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Genes Dev. 11:1381-1395(1997),
```

```
Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
Clegg S., Cobley V.B., Collier R.B., Connor R.B., Corby N.R.,
Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
Bilingron A.G., Frankland J.A., Fraser A., French L., Garner P.,
Grafham D.V., Griffiths C., Griffiths M.N.D., Gvilliam R., Hall R.E.,
Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
Hammond S., Harley J.L., Heath P.D., No., Grilliam R., Johnson C.M., Johnson D.,
Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
Lackacelainho M.H., Leversham M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
Marsh V.L., Martin S.L., McConnachie L.J., McLay K., McChurray A.A.,
Milne S.A., Mistry D., Moore M.J.P., Mullikin J.C., Nickerson T.,
Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
Phillimce B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
Swann R.M., Sycamore N. Taylor R., Tee L., Thomas D.W., Thorpe A.,
Khitchead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
Whitming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ceil. Biol. 20:340-351(2000).

FUNCTION: Makes part of TFIID is a multimeric protein complex that plays a central role in mediating promoter responses to various activators and repressors. Potentiates transcriptional activation by the AF-2S of the retinoic acid, vitamin D3 and thyroid hormone. SUBUNIT: TFIID is composed of TATA binding protein (TBP) and a number of TBP-associated factors (TAPS). Component of the TFTC-HAT complex, at least composed of TAFSL, TAFSL, TAPSL, SUBTISH/SPT3, TAFSLTAPILISO, TAF4/TAFILISO, TAF4/TAFILISO, TAF10, TAFILISO, TAF10, TAFILISO, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, TAF10, 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Davidson I.; "The human TFIID components TAF(II)135 and TAF(II)20 and the yeast SAGA components ADA1 and TAF(II)68 heterodimerize to form histone-like
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      complex
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             "The DNA sequence and comparative analysis of human chromosome 20.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE=97098442; PubMed=8942982;
Tanese N., Saluja D., Vassallo M.F., Chen J.-L., Admon A.;
Tanese Toloning and analysis of two subunits of the human TFIID
complex: hTAFII130 and hTAFII100.";
Proc. Natl. Acad. Sci. U.S.A. 93:13611-13616(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE=20063193; PubMed=10594036;
Gangloff Y.-G., Werten S., Romier C., Carre L., Poch O., Moras D.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IDENTIFICATION IN THE TFTC-HAT COMPLEX WITH TAF5L; TAF6L; TADA3L; SUPT3H; TAF2; TAF5; TRRAP; GCN5L2 AND TAF10.
MEDLINE-99303588; PubMed=10373431;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 870-918 IN COMPLEX WITH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Brand M., Yamamoto K., Staub A., Tora L., "Identification of TATA-binding protein-free TAFII-containing subunits suggests a role in nucleosome acetylation and signal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE OF 105-1083 FROM N.A., AND PARTIAL SEQUENCE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SIMILARITY: BELONGS TO THE TAF2C FAMILY. SIMILARITY: Contains 1 TAFH/NHR1 domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Biol. Chem. 274:18285-18289(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SUBCELLULAR LOCATION: Nuclear.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Nature 414:865-871(2001).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                and TRRAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            transduction
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ص
. .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Rogera
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  pairs.
Mol. C
```

EMBL; Y11354; CAA72189.1; -. EMBL; AL137077; CAC36006.1; -. EMBL; AL109911; CAC22312.2; -.

EMBL; EMBL;

```
------GLH 1089
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 964 KQLKQRAAAIPPIQV-TKVHEPPREDAAPTKPAPPAPPPPQONLQPESDAPQQPGSSPRGK 1022
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1135 V-PYSEHAKAPVGPV-----TMGLP-LPMDPKKLAPFSGVKQE-------Q 1171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         479
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ASGNEVPRGECSGPATVNN 744
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          745 SSDTESIPSPHTEAAKDTGONGPKPP----ATLGADGPP-PGPPTPPRRTSRAPIEPTPA 799
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      76 AAPAEGAPGA-----RPEPPAGRARPGGGGPQRPGPPSP-----RRPLVPA-- 117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  150 -----PEPAPAGPAKPAGPAALAARAGPGPGPGPGPGPGPGKPAGPGAAOTLNGSAAL 202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 203 LNSHHAA---APAVSLVNNGPAALLPLPKPAAPGTVIQTPPFVGAAAPPAAPSPPAAP 259
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                260 APAAPAAPPPPPPPPAATLARPPGHPAGPPTAAPAVPPPA-----AAQNGGSA--GA 309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               310 APAPAPAAGGPA-GVSGQPGPG----AAAAAPAPGVKAESPKRVVQAAPPAAQTLAAS 362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         75
                                                                                         .; TAS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1090 DTARPV----LPRPPTISNP-----PP--LISSAKHPSVLERQIGAISQGMSVQLH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              423 ATLTPTVLAPRLPQPP--QNPTNIQNFQLPPGMVLVRSENGQLLMIPQQALAQ-MQAQAH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         16 EVDEKVVSDLVGSLESQLAASAAHHHLAPRTPEVRAAAAGALGNHVVSGSPAGAAGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       800 SEATGAPTPPPAPPSPSAPPVVPKEEKEETAAAPPVEEGEEQKPPAAEELAVDTGKAE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----GPAPPAAKLRPP-----PEGSAGACAPV-----PAAAAVAAG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     860 EPVKSECTEEAEEGPAK--GKDAEAAEA----TAEGALKAEKKEGGSGRATTAKSSGAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   913 ODSDSSATCSADEVDEAEGGDKNRLLSPRPS----LLTP----TGDPRANASPOKPLDL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1023 SRSPAPPADKEAFAAEAQKLPGDPPCWTSGLPFPVP-----PREVIKASPHA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         363 GPASTAASMVIGPTMQGALPSPAAVPPPAPGTPTGLPKGAAGAVTOSLSRTPTATTSGIR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels 341;
                                                                        GO; GO:0005669; C:transcription factor TFIID complex; TAS.
GO; GO:0016251; F:general RNA polymerase II transcription fac.
GO; GO:0005515; F:protein binding; TAS.
GO; GO:0003713; F:transcription co-activator activity; TAS.
InterPro; IPR007900; TAF4.
InterPro; IPR003894; TAF_hom.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2.5%; Score 329.5; DB 1; Length 1083;
                                                                                                                                                                                                                                    POLY-HIS.
POLY-ALA.
POLY-ALA.
POLY-ALA.
POLY-BRO.
POLY-PRO.
POLY-ALA.
POLY-ALA.
POLY-ARO.
POLY-ARO.
POLY-ARO.
POLY-ARO.
POLY-ASP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----PDPSAFSYAPPGHPLPL------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MW; A6453827572A0752 CRC64;
                                                                                                                                                                                                        protein; 3D-structure
                                                                                                                                                                                                                                                                                                                                                                                               (IN REF. 3).
A -> S (IN REF. 2).
G -> GPG (IN REF. 2).
MISSING (IN REF. 3).
P -> L (IN REF. 3).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ; Pred. No. 0.00044; 94; Mismatches 336;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           701 EDEEMEASGVSGNEEEMVEEAEALH-----
                                                                                                                                                                                                        Transcription regulation; Nuclear
                                                                                                                                                                                                                                                                                                                                                                                                                136 136
185 185
233 264
293 293
1083 AA; 109943 M
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             23.0%;
EMBL; U75308; AAC50901.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity 23.0
Matches 230; Conservative
                                                                                                                                                                       Pfam; PF05236; TAF4; 1.
SMART; SM00549; TAFH; 1.
                                            Genew; HGNC:11537; TAF4.
                                                              MIM; 601796;
                                                                                                                                                                                                                                                                                                                                                                                                                CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1070
                                                                                                                                                                                                                                                                                                                                                                                CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                               CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CONFLICT
                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                      DOMAIN
                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                     DOMAIN
셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð . g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
```

```
1211 GIPSTRVPSDSAITYRGSITHGTPADV-----LYKGTITRIIGEDSPSRLDRGREDSLP 1264
                                                                                                                                                                                                                                                        1373
647
                                                                                                                                                                                                                                                                                                                                                       176
                                                                                                                                                                                                                                                                                                               1374 BAKLLKREGTPPPP----PPSRDLTEAYKTQALGPL-KLKPAHEGLVATVKEAGRSIHE
                                                                                                                                                                                                                                                                                                                                                                                1428 IPREELRHTPEL----PLAPRPLKEGSITQGTPLKYDTGASTTGSKKHDVRSLIGSPGRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M., Murakawa K., Kamakami B., Sugituk Y., Sugano S., Nagahari K., Masuho Y., Nagai K.,
                                                                       QPYLVPFLKRSLPALRQLTPDSAAFIQQSQQQ-----PPPPTSQATTALTAVVLSSSVQR
                                                                                                                                                                                                                                                                                                                                                                                                    596 NFLSTLIKLASS----GKQSTETAANVKELVQNLLDG----KIEAEDFTSRLYR-ELNSSP
                                                                                                                                                                                 1265 KGHVIYEGKKG----HVLSYEGGMSVTQCSKEDGRSSSGPPHETAAPKRTYDMME----
                                                                                                                                                                                                                                                 --GRVGRAISSASIEGLMGRAIPPERHSPHHLKEQHHIRGSITQGIPRSYVEAQEDYLRR
                                                                                                                                                                                                                                                                              703 TAGKTAATVTSÅLOPPVLSLTOPTQ--------vdvGKQGQPTPLVIQQ-----
                                                                                                                                                                                                                                                                                                                                                    -----PPKPGALIRPPQVTLTQT------PMVALRQPHNRIMLT-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE=21456161; PubMed=11572484;
Nagase T., Kikuno R., Ohara O.;
"Prediction of the coding sequences of unidentified human genes. XX
The complete sequences of 60 new cDNA clones from brain which code
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia, Eutheria, Primates, Catarrhini, Hominidae, Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Wambutt R., Heubner D., Mewes H.-W., Weil B., Wiemann S., Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.-!- SIMILARITY: Contains 1 methyl-binding (MBD) domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LSP-----RGQAGPPESLGVPTAQEASVLRGTALGSVPGGSIT-
                                                                                                                                                                                                                                                                                                                                                                                                                                                1484 FPPVHPLDVMA------DAR-----ALERACYEES 1507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DDDINDVASMAGVNLSEESARILATNSELVGTLTRSCKDET 869
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MBD6_HUMAN STANDARD; PRT; 1003 AA. Q96DN6; Q8N3M0; Q8NA81; Q96Q00; 10-0CT-2003 (Rel. 42, Created) 10-0CT-2003 (Rel. 42, Last sequence update) 10-0CT-2003 (Rel. 42, Last annotation update) Methyl-CpG binding domain protein 6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                large proteins.";
DNA Res. 8:179-187(2001).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MBD6 OR KIAA1887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              rissue=Amygdala;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TISSUE-Testis;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TISSUE-Brain;
                                                                                                                                                                                                                                                 1316
                                                                                                                                                                                                                                                                                                                                                                                                                 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 829
                                                                                                                                                                                                                                                                                                                                                744
                                             ठ
                                                                              윱
                                                                                                               à
                                                                                                                                          셤
                                                                                                                                                                               à
                                                                                                                                                                                                            d
                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation—the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1237 VLYKGTITRIIGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYE-----GGMSVTQC 1289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SKEDGRSSSGPPHETAAPKRTYDMMEGRVGRAISSASIEGLMGRAI---PPERHSPHHLK 1346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1347 EQHHIRGSITQGIPRSYVEAQEDYLRREAKLLKREGTPPPPPPPRRDLTEAYKTQALGPLK 1406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LKPAHEGLVATVKEAGRSIHEIPREELRHTPELPLAPRPLKEGSITQGTPLKYDTGASTT 1466
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1467 GSKKHDVRSLIGSP----GRIFPPVH-PLDVMADARALERACYEESLKSRPGTASSSG-G 1520
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1667
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SIARGAPVIVPELGKPROSPLTYEDHGAP-----FAGHLPRGSPVTMREPTPR 1568
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1569 LORGSLSSS---KASQDR-----KLTSTPREIAKSPHSTVPEHH--PHPISP 1610
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1668 Y------LIRGYPDTAALENRQTIINDYITSQQMHHNTATAMAQRADM 1709
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1744 VPPTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPG---GPTHLTK---PTTTSSSERE 1797
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEEDMIKLCNHRRKAVAMATLYRSMETTCSHSSPGEGASPQMFHTVSPGPPSARPPCRVP 152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              153 PTTPLNGG-------PGSLPPEPPS--VSQAFPTLA-GPGG 183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LFPPR--LADPVPSGGSS---SPRFLPRGNAPSPAPPPP---PAISLNAP-SYNWGAALR 234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            235 SSL---VPSDLGSPPAPHASSSPPSDPPLFHCSDALTPPPLPPSNNLPAHPGPASQPPVS 291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1177 QAGPPESLGVPTAQEASVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPAD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RAGGPVATSVPIGWQRCVREGAVLYISPSGTELSSLEQTR------SY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels 370; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 348 AQAPSASHSSSLRPSQRRPRRPPTVFRLLEGRGPQTPRRSRPRAPAPVPQPFSLPEPSQP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     611 YEHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLDAAAAYYL---PRHLAPNPTYPHLYPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       452 OPRHPIQPSLPGTISGSLSSVPGAPAPPAASKAPVVPSPVLQSPSEGLGMGAGPACPLPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               512 LAGGEAFPFPSP-EQGLALSGAGFPGMLGALPLPLSLGQPPPSPLLNHSLFGVLTGGGGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1710 LRG-----LSPRESSLALNYAAGPRGI-----IDLSQVPHLPVL----
                                                                                                                                                                                                                                                                                                                                                                                                                      Length 1003;
                                                                                                                                                                                                                                                                                                                                             D -> G (IN REF. 1; BAB71176).
Q -> QQ (IN REF. 1; BAC04045).
W; 6C8E8693AA6A3BE6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                   Score 325; DB 1; L. Pred. No. 0.00057; 96; Mismatches 449;
                                                                                                                                                                                                                                                                                                                             PRO-RICH
                                                                                                                                                                                                                                     Genew; HGNC:20445; MBD6.
InterPro; IPR001739; Methyl-CpG_bind.
                                                                                                                                                                                                                                                                                                                                                                                 101200 MW;
                                                                                                                                                           EMBL; AK056399; BAB71176.1; -. EMBL; AK093078; BAC04045.1; -.
                                                                                                                                                                                                 EMBL; AB067474; BAB67780.1; -. EMBL; AL834230; CAD38908.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                     22.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Conservative
                                                                                                                                                                                                                                                                                                                            846
637
802
                                                                                                                                                                                                                                                                   Pfam; PF01429; MBD; 1.
SMART; SM00391; MBD; 1.
DOMAIN 17 61
                                                                                                                                                                                                                                                                                                                                                                                 1003 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches 265;
                                                                                                                                                                                                                                                                                                                                                              CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1521
                                                                                                                                                                                                                                                                                                                                             CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     93
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1407
                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                      Local
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            염
```

```
2083
                                                                                                                                                                                                                 ---IITAVEPSKPTVLRSTS 1902
                                                                                                                                                                                                                                                                                                                                                              TSSP-----VRPAATFPPATHCPLGGTLD-----GVYPTLMEP----- 1935
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           --VLLPKEAPRVARPERPRADTGHAFLAK----PPAR--SGLEPASSPSKGSEPRPLVPP 1987
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1988 VSGHATIARTPAKNLAPHHASPDPPAPP--------ASASDPHREKTQSK 2029
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GELRPKQPGPVKLGGEAAHLPHLRPLPESQPSSSPLLQTAPGV-----KGHQRVVTLAQH 2138
                                                                    1798 RDRDRERDRDREREKSILTSTTTVEHAPIWRPGTEQSSGSSGSSGSGGGGSSSRPASHSHA 1857
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     879
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     921
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GGAEPKDPPP-----PGPHSED-----LKVPPGVVRKSRRGRRRKYNPTRN 962
                                                                                                                                                                                                                                                                   571 PPPEPLLPP-----PGGPGP----PLAPGEPEGPSLLVASLLPPPPS----
                                                                                                                                         609 -----DLLPPPSAPPSNLLASFLPLLALGPTAGDG-EGSAEGAGGPSGEPFSGL--
                                                                                                                                                                                                                                                                                                                                                                                                            774 GLQLLPGGG---APPPLSEASSPLACLLQSLQ1PPEQPEAPCLPPESPASALEPEPARPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 831 LSA------LAPPHGSPDPPVPELLTGRGSGKRGRRGGGLRGINGEARPARGRK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2030 PFSIQELELRSL-----GYHGSSYSPEGVEPVSSPSLTHDKGLPXHLEELDKSHLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     880 PGSRREPGRLALKWGTRGGFNGOM-----ERSPRRTHHW---QHNGEL----AE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A. (ISOFORM 2).

TISSUE-Umbilical vein endothelial cells;

MEDLINE=98269042; PubMed=9604001;

MAEDLINE=98269042; PubMed=9604001;

MAEDLINE=98269042; PubMed=9604001;

MEDLINE=98269042; PubMed=9604001;

MEDLINE=98269042; PubMed=9604001;

MEDLINE=98269042; Nimokado K., Miyata T.;

endothelial cells: 12 up-regulated distinct genes including 5 cell growth-related, 3 thrombosis-related, and 4 others.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             002952; 000310; 000498; 099970; 01-FEB-1994 (Rel. 28, Created) 01-FEB-1994 (Rel. 28, Created) 15-MAR-2001 (Rel. 40, Last sequence update) 15-MAR-2004 (Rel. 43, Last annotation update) A-kinase anchor protein 12 (A-kinase anchor protein 250 kDa) (AKAP 250) (Myasthenia gravis autoantigen gravin).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Nauert J.B., Klauck T.M., Langeberg L.K., Scott J.D.; "Gravin, an autoantigen recognized by serum from myasthenia gravis patients, is a kinase scaffold protein."; Curr. Biol. 7:52-62(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TISSUB-Umbilical vein endothelial cells;
Bowditch R.D., Ginsberg M.H.;
"Sequence of gravin cDNA isolated from a human umbilical vein
endothelial cell library."
Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2139 ISEVITQDYTRHHPQQLSAPLPAPLYSFPGASCPVLDLRR 2178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -SNSSRODITLEPSPTARAAVPLPPRARPGR--PAKNKRR 999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT; 1781 AA
                                                                                                                                                                                                                 1858 HQHSPISPRTQDALQQRPSVLH-NTGMKG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE=97153077; PubMed=9000000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A. (ISOFORM 1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Biochem. 123:1119-1126(1998)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE OF 43-1781 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AKAP12 OR AKAP250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TISSUE=Heart
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1936
                                                                                                                                                                                                                                                                                                                                                                  1903
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2084
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  922
                                                                                                                                                                                                                                                                                          657
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   963
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AK12_HUMAN
AK12_HUMAN
AK12_H
AK12_H
AK12_H
DT 16-OCT
DT 16-OCT
DT 16-OCT
DE 250) (f
EN AKAP12_G
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY
CO ENKARY

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 83
                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                    ò
                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                                                                                                                                                            용
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ઠે
```

```
SWISS-PROT entry is copyright. It is produced through a collaboration een the Swiss Institute of Bioinformatics and the EMBL outstation -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      noved. Usage by and for commercial (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                           Note-No experimental confirmation available;
--- TISSUE SPECIFICITY: Expressed in endothelial cells, cultured fibroblasts and osteoaarcoma, but not in platelets, leukocytes, monocytic cell lines or peripherical blood cells.
--- INDUCTION: Activated by lysophosphatidylcholine (lysoPC).
--- DOMAIN: Polybasic regions located between residues 265 and 556 are involved in binding PKC.
--- DISEASE: ANTIBOLIES TO THE C-TERMINAL OF GRAVIN CAN BE PRODUCED BY PATIENTS WITH MYASTHENIA GRAVIS (MG).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        between the Swiss Institute of Bioinformatics and the EMBL outstatistic the Buropean Bioinformatics Institute. There are no restrictions on use by non-profit institutions as long as its content is in modified and this statement is not removed. Usage by and for commer entities requires a license agreement (See http://www.isb-sib.ch/annou
                                                                                                                                                         U
                                                                                                                                                                                     -!- SUBUNIT: Binds to dimeric RII-alpha regulatory subunit of PKC.
                                                                                                                                                         and protein kinase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /FTId=VSP 004110.
EEEVIVTE -> MLGTITIT (in isoform 2)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pint, 020005737; C:cytoplasm; TAS.

GO; GO:00050737; C:cytoplasm; TAS.

GO; GO:0005079; F:protein kinase A anchoring activity; TAS.

GO; GO:0007186; P:G-protein coupled receptor protein signalin.

InterPro; IPR001573; PkinA anch.

Pfam: PF03832: PkinA anch: 3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (PROBABLE)
             SEQUENCE OF 1477-1781 FROM N.A.
TISSUB-Umbilical vein endothelial cells;
MEDLINE=92395179; PubMed=1522245;
Gordon T., Grove B., Loftus J.C., O'Toole T., McMillan R.,
Lindstrom J., Ginsberg M.H.;
"Molecular cloning and preliminary characterization of a novel
                                                                                                       cytoplasmic antigen recognized by myasthenia gravis sera.";
J. Clin. Invest. 90:992-999(1992).
-!- FUNCTION: Antoring protein that mediates the subcellular compartmentation of protein kinase A (PKA) and protein kir
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | ZETJG=VSP 004111.
| E -> K (IN REF. 2).
| NRN -> TPEI (IN REF. 2 AND 3)
| Q -> K (IN REF. 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BA813937379FAC0F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       > E (IN REF. 2 AND 3).

> G (IN REF. 2 AND 3).

> S (IN REF. 2 AND 3).

> A (IN REF. 2 AND 3).

> EE (IN REF. 3 AND 4).
                                                                                                                                                                                                                                                                                                                  IsoId=Q02952-2; Sequence=VSP 004110, VSP 004111;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              INVOLVED IN PKC-BINDING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DING (PROBABLE)
(in isoform 2)
                                                                                                                                                                                                                                                Event=Alternative splicing; Named isoforms=2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (IN REF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RII-BINDING
Missing (in
                                                                                                                                                                                                                                                                                   IsoId=Q02952-1; Sequence=Displayed;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AKAP 1.
AKAP 2.
AKAP 3.
POLY-GLU.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  splicing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ^ ^
≅ ≅
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; AF001504; AAB58938.1; -. EMBL; AB003476; BAA19927.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; U81607; AAC51366.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Antigen; Repeat; Alternative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; M96322; AAA35931.1; -. Genew; HGNC:370; AKAP12.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AA; 191439
                                                                                                                                                                                                                                     ALTERNATIVE PRODUCTS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            101
556
1553
98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           106
                                                                                                                                                                                                                      cytoskeleton.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           694
867
986
1530
1581
1601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      604698;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            This
```

ch 1 Similarity 18.7%; Pred. No. 0.00096; 417; Conservative 273; Mismatches 727; Indels 810; Gaps 102;	75 GNERSQELHLRPESHSYLPELGKSEMEFIESKRPRLELLPDPLL 118	38 GQEENRNIEQIPSSESNLEELTQPTESQANDIGPKKVFKFVGFKFTVKKDKTE-KPDTVQ 196	19 RPSPLLATGOPAGSEDLTKDRSLTGKLEPVSPPSPPTDPELELVPPRLSKEELIQNMDR 178	97 LLTVKKDEGEGAAGAGDHQDPSL-GAGEAASKESEPKQSTEKPEETLKREQ 246	σ		L)	-	OWKOKFCORYDOLMEALEKKVERIENNPRRRAKESKVREYYEKOFPEIRKOR	29 QEPEKVDTEEDGKAEVASEKITASEQAHPQEPAESAHEPRISAEYEKVELPSEEQVS- 385	49 RMQSRVGQRGSGLSMSAARSEHEVSEIIDGLSEQENLEKQMRQLAVIPP 397	36GSQGPSEEKPAPLATEVFDEKIEVUQEEVVAEVHVSTVEERTEEQKTEV 434	MLYDADQQRIKFINMNG	: : : : : : : : : : : : : :	FMOHPKNFG	: : : :: : :: : :: : : :: : : : : : : : : : : : : : : : : : :	500000000000 6	16 HTQ	99 DEKEAVASKGRKTANSQGRRKGRITRSMANBANSEAITPQOSAE 603	O EAEEGATSDGEKKREGVIPWASFKKMVTPKKRVRRPSESDKEDELDKVKSAT 641	4 LASMELNESSRWTEEEMETAKKGLLEHGRNWSALARMVGSKTVSQCKN 651		FYFNYKKRONLDEILQQHKLKMEKER		1 EDEEMEASGVSGNEEEMVEEAEALHASGNEVPRGECSGPATVNNSSDT-ESIPSPHTEAA 759	8 KDKETGIDGILAGSQEHDPGQGSSSPEGAGSPTEGEGVSTWESFKRLVTPRK 769	0 KDTGQNGPKPPATLGADGPPPGPPTPPRRTSRAPIR- 795		PTPASEATGAPTPPPAP	O AGPTGANEDDSDVPAVVPLSEYDAVEREKMEAQQAQKGAEQPEQRAAT 877	DELAVOTGKAE 869		0AKGKDABAABATABGALKABKKGGSGRATTAKSSGAPQDSDSS 918	8 REVIAEEEPPTVTEPLPENREARGDTVVSEAELTPEAVTAAETAGPLGS 986	9 ATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRANASPQKPLDLKQLKQRAAAI 973
Query Match Best Local Sim Matches 417;		138 GC	119 RF	197 LL		247 SH		305 GW	293 QW			386	398 ML	435	444 FM	: 487 VL		546 HT		590 EA	604 LA	642 LS	652 FY		701 EDI	718 KD	760 KDJ	770 KSF	1 964	830 AGE	850 ELA	878 EVS	870	938 REV	919 ATC
Query Best Match	ζ	Q .	ò	QQ	ò	Db	ò	qq	ζ	οp	ζ	Ор	ò	අු	ò	qq	ò	QQ	ò	O p	ò	qq	δλ	QQ	ò	q	ò	ΟD	ò	Op	X O	ОЪ	ò	ą	δ

qa	
<i>\</i> 0	974 PPIQVTKVHEPPREDAAPTKPAPPAPPPQNLQPESDAPQQPGSSPRGKSRSPAPPADKE 1033
QQ	1042 LQAVAEKVKEESQLPGTGGPEDVLQPVQRAEAERPEEQAEASGLKKE 1088
که ا د	DPSAFSYAPPGHPLPLGLHD 109
ò	1 TARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSV
Op	QVTESIESSE 113
δλ	GQAGPPESLGVPTAQEASVLRGTALGSVPGGSITK 12
ф	
ò	YKGTITRIIGEDS
DÞ	1189 EIVEIHEENEVASGTQSGGTEAEAVPAQKERPPASSFVFQ 1229
ð í	1 BCKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHI
an i	9 3 3
දු ද	PRSYVEAQEDYLRRE 137
3 (
è é	5AKLLKREGTPPPPPPRDLTEAYKTOALGPLKLKPAHEGLVATVKEAGRSIHE 142
g	4,
දු දු	PPV 148
}	OR HDI DIMADADAT BOANGEDES CANGED TO MACCOCCATACA TITE TO SECOND TO THE TAXACA
3 20	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
'n	KLTSTPREIAKSPHSTVPEH 160
qq	 -VP 147
δ	1604 HPHPISPYEHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPH 1663
đ	
ò	LENROTI INDY:
Db	1499LEG-EKTISLKWKSDEVDEQVACQEVKVSVAIEDLEPENGILELETKSSKLVQN 1551
δ	1724 YAAGPRGIIDLSQVPHLPVLVPPTPGTPATAMDRLAYL-PTAPQPFSSRHSSSPLSPGGP 1782
Dp	15521IQ1AVDQFVRTEETATEMLTSELQTQ 1578
'n	1783 THLTKPTTTSSSERERDRDREREREKSILTSTTTVEHAPIWRPGTEOSSGSSG 1842
Ор	DAGOETEI
à	1843 GGGGSSSRPASHSHAMOHSPISPRTQDALQQRPSVLHNTGMKGIITAVEPSKPTVLRSTS 1902
οp	1626HSDISKDMSEASEKTMTVEVE 1646
δλ	1903 ISSPVRPAATFPPATHCPLGGTLDGVYPTLMEPVLLPKEAPRVARPERPRADTGH 1957
qq	1647DDGLEEVVLPSEEEGGGGAGTKSVPEDDGH 1680
λõ	1958 AFLAKPARSGLEPASSPSKGSEPRPLVPPVSGHATIARTPAKNLAPHHASPDPPAPPAS 2017

307 BALEKKVERIENNPRRRAKESKVREYYEKQPEIRKQRELQERMQSRVGQRGSGLSMSAA 366	
8 8 8 8 8	8 4 8 4 8 4 8 4 8 4 8 4 8 4 8 8 8 8 8 8
Db 1681 ALLAERIEKSLVEP-KEDEKGDDVDDPENQNSALADTDASG-GLTKESPDTNGP 1732 Qy 2018 ASDPHREKTQSKPFSIQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDKGL-PKHLEE 2076	PERSULT 84

us-09-522-753-5.rgp

8 8 8 8 8 8	D REC	P P P P P P P P P P P P P P P P P P P	R R R R R R R R R R R R R R R R R R R	E488888888888888	38888888888888888
1244 TRIIGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGCMSVTQCSKEDGRSSSGPPHE 1303		2426 LKTLGQVRVVAPLLYCDGHSEPF	2530 YFNAAGASKWPTDVLNPSFYEHEDPPLPVGYQLPPNPRNVQELFSGFPPRVGHG 2583 1750TPATAMDRLAYL		1971 PASSPEKGSEPRDLVPPVSGHATIARTPAKNILAPHHASPDPPAP 2014
6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	6 6 6 6	8 6 8 6 8 6	9	60 60 60 60 60 60 60 60 60 60 60 60 60 6	6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6

```
2972 MHSG-----SEPGAR-PEVRLSQYRHAGPQTYTVRKEAPPSAASQLPKNPKCKDSMY 3022
                                                                                                                2301
                                                                                                                                                                                 3050
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE-93203161; PubMed=1284248;

RX Yamaguchi N., Kimura S., McBride O.W., Hori H., Yamada Y.,

RA Yamaguchi N., Ximura S., McBride O.W., Hori H., Yamada Y.,

RA Kanamori T., Yamakoshi H., Nagai Y.; McBride O.W., Hori H., Yamada Y.,

RA Kanamori T., Yamakoshi H., Nagai Y.; McBride Collagenous domains and partial characterization of a novel collagenous foath, alpha 1(XVI), consisting of repetitive collagenous domains and Cysteine-containing non-collagenous segments.";

RI "Molecular cloning anon-collagenous segments.";

RI "Molecular cloning non-collagenous segments.";

C --- FUNCTION: The numerous interruptions in the triple helix may make contain molecule either alastic or flexible.

CC --- FUNCTION: The numerous interruptions in the triple helix may make contain, a membranous tissue lining the amniotic cavity. Within the complex network of reticular fibers. Also located to a fibroblast complex network of reticular fibers. Also located to a fibroblast collagen.

CC --- DEVELOPPRIVILE STAGE: Transiently elevated expression during
                                                                                                                                                                                                                                              2361
                                                                                                                                                                                                                                                                          This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-aib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                               2242 YPLLYRDGEQTEPSRMGSKSPGNTSQPPAFFSKLTESNSAMVKSKKQBINKKLNTHNRNE
                                                                                                                                                                                                                             2302 PEYNISQPGTEIFNMPAITGTGLMTYRSQAVQEHASTNMGLEAIIRKALMGKYDQWEESP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
MEDLINE-9233539; PubMed=1631157;
Pan T.C., Zhang R.Z., Mattei M.-G., Timpl R., Chu M.-L.;
"Cloning and chromosomal location of human alpha 1 (XVI) collagen.";
Proc. Natl. Acad. Sci. U.S.A. 89:6565-6569 (1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               gestation, and decrease at term.

DOMAIN: This sequence defines eighteen different domains, nine triple-helical domains (COL9 to COL1) and ten nontriple-helical
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Buteleostomi,
Mammalia; Eutheria, Primates, Catarrhini, Hominidae, Homo.
NCBI_TaxID=9606;
                                                                                                                                               Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT; 1603 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          chain precursor
                                                                                                                                                                                                                                                                                                                                                                                                     3082 STSGDLVVPSGSPSSLSTAAP 3102
                                                                                                                                                                                                                                                                                                                                                               2362 PLSANAFNPLNASASLPAAMP 2382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-FEB-1995 (Rel. 31, Created)
01-FEB-1995 (Rel. 31, Last seq
10-OCT-2003 (Rel. 42, Last ann
Collagen alpha 1(XVI) chain pr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE OF 418-1603 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; M92642; AAA58427.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (Human)
                                                                                                                                                                3023 YP-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HUMAN
                                                                                                                                                                                                                                                                                                  3051
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        007092
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ALF HUMAN
```

GEKGDGGIKGVPGKPGRDAPGEICVIGPKGQKGDPGFVGP HEPPREDAAPTKPAPPAPPPQNLQPESDAPQQPGSSPRGKSRSPAPPADKEAFAAEAQK	442BGLAGEPGPFGLPGPPGIGLPGTPG-DPGGPPGPKGDKGSSGIPGKE 487 1042 LPGDPPCWTSGLPFPVPPR-EVIKASPHAPDPSAFSYAPPGHPLPLGLHDTARPVLPRPP 1100 448	TISNPPPLISSAKHPSVLERQIGAISQGMSVOLHVPYSEHAKAPVGPVTMGLP-LP TISNPPLISSAKHPSVLERQIGAISQGMSVOLHVPYSEHAKAPVGPVTMGLP-LP TISNPPLISSAKFSVLERQIGAISQGMSVOLHVPYSETGASGDVGSPGFGLPGLP TRARGDPGIQAIXGEKGEPCLSCSSVVGAQ-HLVSSTGASGDVGSPGFGLPGLP	1156 MDPKKLAPFSGVKQEQLSPRQQAGPPESLGVPTAQEASVLRGTALGSVPGGSITKGIPST 1215 589GRAGVPGLKGRK-GNPCBAGFPGSPGPPGPVGPAGIKGAKGEPCE 632	Σ Ω	HVLSYEGGMSVTQCSKEDG : : KAGERGLKGQ	1336 PPERHSPHHLKEQHHIRGSITQGIPRSYVEAQEDYLRREAKLLKREGTP- 1384	PPPPPSRDLTBAYKTQALGPLKLKPAHEGLVATVKEAGRSIHEIPREELRHTPE	QTAHAdd	811 GPPGPTGEKG-AQGSPGVKGATGPVGPPGASVSGPPGRDGQQGQT 854 1492 VMADARALERACYEESLKSRPGTASSSGGSIARGAPVIVPELGKPRQSPLTYEDHGAPPA 1551	855GLRGTPGEKGPRGEKGEPGECSCPSGGDLIFSGMPGA 891	PGLWMGSSWQPGPQGPPGIPGPPGPPGVPGVPGNNGLPGQPGLTAELGSLPIE	1596 PHSTVPEHHPHPISPYEHLIRGVSGVDLYRSHIPLAFDPTSIPRGIPLDAAAAYYLPR 1653	1954 HLAFWRITEHINFFILL	ABEARGDNSEGDPGCVGSPGLPGQRGEEGPPGRRGSPGPPG		LEISTITYERHEITER FOLESSOGSOCGOGGSSARASARSARSARSARSARSARATOR	LSKRIQ	1909PAATFPPATHCPLGGTLDGVYPTLMEPVLLPKEAPRVARPERPRADTGHA 1958
ବ ୪ ବ	8 & E	3 & 8	<i>ኤ</i> ଶ	ζζ Dp	oy Oy	<i>አ</i>	l & 1	8 &	d V	a 8	3 8	ያ ያ	8 8 8	8 8 8	S 8 5	g - 6	<u>\$</u> 8	B &
EMBL; S57132; AAB25797.1; PIR; S23810; S23810. Genew; HGNC:2193; COL16A1. MIM; 120326; -:	GO; GO:0005557; C:collagen type XVI; TAS. GO; GO:0007565; P:pregnancy; TAS. InterPro; IPR008160; Collagen. InterPro; IPR0088165; ConAlike_lec_gl.	Pfam; PP01391; Collagen; 18. Pfam; PP01391; Collagen; 18. SMART; SM00210; TSPN; 1. Extracellular matrix: Connective tissue; Collagen; Hydroxylation;	Repeat; Signal. SIGNAL 1 21 POTENTIAL. CHAIN 22 1603 COLLAGEN ALPHA 1(XVI) CHAIN. DOMAIN 50 231 TSP N-TERMINAL.	232 374 375 505 506 520	DOMAIN 521 554 TRIPLE-HELICAL REGION 8 (COL8) DOMAIN 555 571 NONHELICAL REGION 8 (NC8). DOMAIN 572 630 TRIPLE-HELICAL REGION 7 (COL7)	WITH 1 IMPERFECTION. DOWAIN 631 651 NONHELLCAL REGION 7 (NC7). DOWAIN 652 722 TRIBLE-HELLCAL REGION 6 (COL6) WITH 1 IMPERFECTION	723 737 738 875	87 886 887 938 939 972	DOMAIN 973 987 TRIPLE-HELICAL REGION 3 (COL3). DOMAIN 988 1010 NONHELICAL REGION 3 (NC3). DOMAIN 1011 1432 TRIPLE-HELICAL REGION 2 (COL2) MITTH 2 IMPREPREMENTONS 1011	DOMAIN 1433 1471 NONHELICAL REGION 2 (NC2). DOMAIN 1472 1577 TRIPLE-HELICAL REGION 1 (COL1) MITTLE J TRIPLE-HELICAL REGION 1 (COL1)	1578 1603 NONHEL 418 420 RDA -> 537 537 R -> P	CONFLICT 1160 1160 T -> P (IN REF. 2). CONFLICT 1163 1163 T -> P (IN REF. 2). CONFLICT 1165 1165 S -> P (IN REF. 2). SEQUENCE 1603 AA; 157692 MW; E27D9AlD4E598A37 CRC64;	ery Match st Local Similarity 21.4%; Pred. No. 0.00098; tches 376; Conservative 143; Mismatches 632; Indels 609; Gaps 99;	751 IPSPHTEAAKOTGQN-GPKPPATLGADGPPRG	782PPTPPRRTSRAPIEPTPASEATGAPTPPPAPPSPSAPPVV 822 	823 PKEEKBEETAAAPPVEBGEEQKPPAAEELAVDTGKAEEPVKSECTEEAEBGPAK 876	877 GKDAEAAEATAEGALKAEKKEGGSGRATTAKSSGAPQDSDSSATCSADEVDEAEG 931 358 GNDCYRISPDAPLQCAEGP-KGEKGESGALGPSGLPGSTGEKGQK 401	932 GDKORLLSPRPSLLTPTGDPRANASPOKPLDLKQLKQRAAAIPPIQVTKV 981

6 8 6 8 6 8 6

```
1251 GLPGPKGDCGKPGPPGSTGRPGAEGEPGAMGPQGRPGPPGHVGPPGPPGQPGPAGISAVG 1310
                                                                                                            1357
                                                                                                                                                                                 1404
                                                                                                                                        GYHGSSYSPEGVEPVSPVSSPSLTHDKGLPKHLEEL-----DKSHLEGELRPKQPGPVK 2095
                                                                                                                                                                                                                    2154
                                                                                                                                                                                                                                           1405 -AGERGH------PGAPGPSGSPGL---PGVPGSMGDMVNYDBIKRFIRQEIIKMFDER 1453
                                                                                                                                                                                                                                                                                                                                                                                     1500 REGRQGLPGVRGLPGTKGEKGDIGIGLAGENGLP--GPPGPQGPPGYGKMGATGPM---- 1553
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                         2096 IGGEAAHLPHLRPLPESO-PSSSPLLQTAPGVKGHQRVVTLAQHISEVITQDYTRHHPQQ
                                                                                                                                                                                                                                                                                                                                                   2204 K--RSPEPNKTSVLG------GGEDGIEPVSPPEGMTEPGHSR-SAVYPLLYRD
                                                                   ----VSGHATIARTPAKNLAPHHASPDPPAPPASASDPHREKTQSKPFSIQELELRSL
                                                                                                                                                                                                                                                                                        9----
                                                                                                                                                                                                                                                                                                                    1454 MAYYTSRMQFPWEMAAAPG-----RPG-----PPGKDGAPGRPGAPGLPGQIG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE-91021039; PubMed=2171211;
WICEK C., Kozmik Z., Paces V., Schirm S., Schwyzer M.;
"Pseudorabies virus immediate-early gene overlaps with an oppositely oriented open reading frame: characterization of their promoter and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VIROLOGY 179:365-377(1990).

-!- FUNCTION: THIS IE PROTEIN IS A MULTIFUNCTIONAL PROTEIN CAPABLE OF MIGHATING TO THE NUCLEUS, BINDING TO DNA, TRANS-ACTIVATING OTHER VIRAL GENES, AND AUTOREGULATING ITS OWN SYNTHESIS.
-!- SUBCELLULAR LOCATION: Nucleus of infected cells.
-!- PTW: A long stretch of serine residues may be a major site of phosphorylation.
-!- SIMILARITY: BELONGS TO THE ICP4/IE140/IE180 FAMILY.
                                                                                                                                                                                                                                                                                LS----APLPAPLYSFPGASCPVLDLRRPPSDLYLPPDHGAPAR-GSPHSEG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pseudorabies virus (strain Kaplan) (PRV).
Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
Alphaherpesvirinae; Varicellovirus.
Notel TaxID=33703;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-FEB-1994 (Rel. 28, Created)
01-FEB-1994 (Rel. 28, Last sequence update)
01-FEB-1994 (Rel. 28, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1446 AA
FLAKPPARSGL -- EPASSPSKGSEPRP-
                                                                                                                                                                                                                                                                                                                                                                                                                            GEOTEPSRMGSKSPGNTSQP 2268
                                                                                                                                                                                                                                                                                                                                                                                                                                                            1554 GOGGIPGIPG--PPGPMGOP 1571
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Immediate-early protein IE180.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRVKA
                                                                   1988
                                                                                                                                          2042
                                                                                                                                                                                                                                                                                    2155
                                                                                                                                                                                                                                                                                                                                                                                                                            2249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    IE18 PR
P33479;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IE18
                                                                                             ద
                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                    à
                                                                                                                                                                                                                                     d
                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
```

PIR; A45344; A45344. InterPro; IPR005205; Herpes ICP4_C. InterPro; IPR005206; Herpes_ICP4_N. Pfam; PF03885; Herpes_ICP4_C; I. Pfam; PF03884; Herpes_ICP4_N; I.

EMBL; M34651; AAA47470.1; -.

```
1008 ESDAPQQPGSSPRGKSRSPA-PPADKEAFAAEAQKLPGDPPCWTSGLPFPVPPREVIKAS 1066
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1290
                                                                                                                                                              745
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    .067 PHAPDPSAFSYAPPGHPLPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLERQIGAIS 1126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1407
                                                                                                                                                                                          128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               KEDGRSSSGPPHETAAPKRTYDMMEGRVGRAISSASI---EGLMGRAIPPERHSPHHLKE 1347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1451
                                                                                                                                                                                                                        782
                                                                                                                                                                                                                                                    183
                                                                                                                                                                                                                                                                                 823
                                                                                                                                                                                                                                                                                                             243
                                                                                                                                                                                                                                                                                                                                        887
                                                                                                                                                                                                                                                                                                                                                                                                                           304 TVARGPEKDEDGLGLAGDGGAPPORQPRRRRAGEGALRRGRGFSSSSSSSSSSDLSPARS 363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TQGTPLKYDTGASTTGSKKHDVRSLIGSPGRTFPPVHP----LDVMADARALERA---- 1502
                                                                                                                                                                                                                                                                                                                                                                       303
                                                                                                                                                                                                                                                                                                                                                                                                                                                               947
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ------LTPLGEPWP------GSDPPADGRVRYGGAGDSREGLWD 510
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            622
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         723
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EEVTAALLALRDAIPGAG-----PAERRQAADSVALVARTVAPLVRYSVDGARAREAAW 777
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       778 TYAAALFAPANVAAARLAEAAARPGPAEPAPGLPPLWPEQPGLVVPAPAPAAGAPSGLP 837
                                                                                                                                                                                                                                   EHGHLVLGPRSRAG----SGPRPPTPAALAAAEAGAPGGPGRSSPSAASPASSSGSPGP
                                                                                                                                                                                                                                                                          783 PTPPRRTSRA---PI-EPTPASE-ATGAPTPPPA------PPSPSAPPPVVP-
                                                                                                                                                                                                                                                                                                       184 SAAPRRWSPARGDPVGEPGPARRPRTPAPPAQPAAVAAAPARRGPASPAAGPVSAPG
                                                                                                                                                        686 APAAASEEAAFPPVVEDEEMEASGVSGNEEEMVEEAEALHASGNEVPRGECSGPATVNNS
                                                                                                                                                                                  AAAAGATRPPRPPSAQQQQPRRG-SGEIVVLDDEDEEEDEPGS--PAAGSPGRALHQGS
                                                                                                                                                                                                                                                                                                                                                                 244 GGGAPSAGGDRGRHHHQHREPLLDEPAAARRLDPRPLGARSPVSSNPNSNSNSTTTVAVE
                                                                                                                                                                                                                                                                                                                                                                                                  --PAKGKDAEAAEATA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  888 EGALKAEKKEGGSGRATTAKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               948 TGDPRANASPQKPLDLKQLKQRAAAIPPIQVTKVHEPPREDAAPTKPAPPAPPPQNLQP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               408 -GAPLARAGP-----PPSPPAPAPRPSASSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1127 QGMSVQLHVPYSEHAKAPVGPVTMGLPLPMDPKKLAP-----FSGVKQEQLS----PRG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          QAGPPESLGVPTAQEASVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPAD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1237 VLYKGTITRIIGEDSPSRLDRGRE----DSLPKGH--VIYEGKKGHVLSYEGGMSVTQCS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----PAPAR-----APAALGPACL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         435 SSSAAASPAPAP----EPARPPRRKRRSTNNHLSLMADGPPPTDG---PL----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RVRAAFAAAQP-----CPLRL-----CRVPRELADACVLACRGVLERLLP----CPLRL--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        QHHIRGSITQGIPRSYVEAQEDYLRREAKLLKREGTPPPPPPPSRDLTEAYKTQALGPLKL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                .408 KPAHEGLVA---TVKEAGRSIHEIPREBLRHTPELPLAPR---PL-----KEGSI
                                                                                                                               582;
            factor;
                                                                                              2.4%; Score 323; DB 1; Length 14 ilarity 20.5%; Pred. No. 0.00092; Conservative 158; Mismatches 590; Indels
                                                                 1446 AA; 148640 MW; 81F43A3DE3DDA068 CRC64;
                                                                                                                                                                                                                  746 SDTESIPSPHTEAKDTGONGPKPP-----ATLGADGPP----
          Trans-acting
                    Phosphorylation, Nuclear protein. 147 354 POLY-SER. POLY-SER.
                                                                                                                                                                                                                                                                                                                                                                                                T------GKAEE----PVKSECTEEAEEG
          regulation;
       protein; Transcription
                                                                                                               Similarity
Early protein, DNA-binding, Phos
                                                                                                                          Matches 344;
                                                                   SEQUENCE
                                                                                                                                                                                                                                              129
                                                                                                                                                                                                                                                                                                                                                                                                855
                                                                                              Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                      824
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             511
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       567
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          699
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             724
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      709
                                                   DOMAIN
       WAFFS
                                                                                                                                                        à
                                                                                                                                                                              g
                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                    엄
                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

```
Genet. 14:33-41(1996).
                                                                                                                                                                                                 HGNC:13013; MYST3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1414
1597
1704
1977
1547
                                                                                                                                                                                                                                                                                                                                                                                                                                     026
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2004 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Local Similarity
tes 282; Conserv
                                                                                                                                                                                                 Genew; HGNC:1301
MIM; 601408; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1411
1593
1643
1897
1546
                                                                                                                                                                                                                                                                                                                                                     Nuclear protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             861
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
ZN FING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                            FING
                                                                                                                                                                                                                                                                                                                                                                   FING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Loc
Matches
8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                         1951
                                                                                                                                                                                                                                                                                                                                                                                                        LDAAAAYYLP----RHLAPNPTYPHLYPPYLIRGYPDTAALENRQTIINDYITSQQMHHN 1698
                                                                                                                                                                                                                                                                                                                                                                                                                                                            - FDAGAAFGHRAANRWGLGAPLRPVFVSCGRRGLAELRGPEGLPAELRAFCAAALLEPDA 1343
                                                                                                                           DRL------ETHTPR--P 988
                                                                                                     RKLTSTPREIAKSPHSTVPEHHPHPISPYEHLLRGVSGVDLYRSHIPLAFDPTSIPRGIP
                                                                                                                                                                                                                                                                                                                                                                                         KPTVLRSTSSPVRPAATFPPATHCPLGGTLDGVYPTL--MEPVLLPKEAPRVARPERP
                                                                                                                                                                                                                                                                                                                                                                                                                                  SAGAT-----
           ----SGGSIARGAPVIVPELG----KPRQSPLT---
                                                                           DGEEDEG--LSGAALRGDGHGHRDDEEDRGPRRKRRSLGLGPAPDPAPALLSSSSSED
                                                                                                                                                                        SEAALAAYCPPEVARALVDQEVFPELWRPALTF---DPAALA------HIAARR----
                                                                                                                                                                                                 TATAMAQRADMLRGLSPRESSLALNYAAGPRGIIDLSQVPHLPVLVPPTPGTPATAMD-R
                                                                                                                                                                                                                       -GAPLRRRAAWMRQIADPED-----VRVVVLYDPLPHEELCAEPAEGAPRPAWDPR
                                                                                                                                                                                                                                              LAYLPTAPQPFSSR-----HSSSPLSPGGPT--HLTKPTTTSSSERE------
                                                                                                                                                                                                                                                                    RGGLSALLAAFAHRLCTPDSHAWAGNWTGRPDIGRLNAQGVLLLSARDLGFAGAVEYLCS
                                                                                                                                                                                                                                                                                             -RDRDRERDRDREREKSILTSTTTVEHAPIWRPGTEQSSGSSGSSGGGGGSSSRPASHSH
                                                                                                                                                                                                                                                                                                                -----LIVLDTIEDWPADGPA--------VGDYH
                                                                                                                                                                                                                                                                                                                                          1857 AHQHSPISPRTQDAL----QQRPSVLHNTGMKG--IITAVEPS------
                                                                                                                                                                                                                                                                                                                                                         ----FAPFETSVRAAAGGAVETH
                               GSGPSSPASTKSSSSTKSSSTKSGLSGSSGYASSPAAGPDPAPERRKKKRRAPGARRPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GHATIARTPAKNLAPHHASPDPPAPPASASDPHREKTQSKPFSIQELELRSLG----Y
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Borrow J., Stanton V.P. Jr., Andresen J.M., Becher R., Behm F.G., Chaganti R.S.K., Civin C.I., Disteche C., Dube I., Frischauf A.M., Horsman D., Mitelman F., Vollnia S., Watmore A.E., Housman D.E.; "The translocation t (8;16) (pilipl3) of acute myeloid leukaemia fuses a putative acetyltransferase to the CREB-binding protein.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2044 HGSSYSPEGVEPVSPVSSPSLTHDKGLPKHLEELDKSHLEGELRPKQP-GPVKL 2096
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1387 RPAGASGAGAGPGEDGDSVEIVGVRG-------GDGRPRGPLGPIKV 1426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      15-JUL-1998 (Rel. 36, Created)
15-JUL-1998 (Rel. 36, Last sequence update)
10-OCT-2003 (Rel. 42, Last annotation update)
WYST histone acetyltransferase 3 (Runt-related transcription factor binding protein 2) (Monocytic leukemia zinc finger protein) (Zinc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Bukaryota, Metazoa, Chordata, Craniata, Vertebrata, Buteleostomi,
Mammalia, Butheria, Primates, Catarrhini, Hominidae, Homo.
                                                       --YEDHGAPFAGHLPRGSPVTMREPT----PRLQEGSL----
                                                                                                                                                                                                                                                                                                                                                                                                                                      -----SKGSEPRP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2004 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1344 EAAPLVLTPG-----AVAAAGAPPAVLWD-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                     RADTGHAFLAKPPARSGLEPASSP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 finger protein 220).
MYST3 OR RUNXBP2 OR ZNF220 OR MOZ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
MEDLINE=96376968; Pubmed=8782817;
           ----CYEESLKSRPGTASS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                   RLGAARRR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MYS3 HUMAN
                                                                              868
                                                                                                                           926
                                                                                                                                                  1643
                                                                                                                                                                          686
                                                                                                                                                                                                 1699
                                                                                                                                                                                                                       1034
                                                                                                                                                                                                                                              1758
                                                                                                                                                                                                                                                                     1084
                                                                                                                                                                                                                                                                                                                                                                 1173
                                                                                                                                                                                                                                                                                                                                                                                         1894
                                                                                                                                                                                                                                                                                                                                                                                                               1233
                                                                                                                                                                                                                                                                                                                                                                                                                                      1952
           1503
                                 838
                                                       1543
                                                                                                                                                                                                                                                                                                                                                                                                                                                             1285
                                                                                                      1583
                                                                                                                                                                                                                                                                                             1798
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MYST
                             8
                                                                                                                                                                      셤
                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ઠ
                                                       ઠે
                                                                        g
                                                                                                   ઠે
                                                                                                                          셤
                                                                                                                                                                                                ઠે
                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                용
                                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                       ઠે
                                                                                                                                                                                                                                                                                                                                                                                                            ద
                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      용
                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ઠ
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstationthe European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce) or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      62;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               921 VASEEQPSQDGKPDLPK-RRLSEGVEP-----WRGQLKKSPEAL-KCRLTEGSERLP 970
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           801 BPQCQERELEISVGKSVSHENKEQDSYSVESEKKPEVMAPVSSTRLSKQVLPHDSLPANS 860
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                125 ATG--OPA--GSEDLTKORSLTGKLEPVSPPSPPHTDPELELVPPRLSKEELIQNMDRVD 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             241 AHRILEGLGPQVELPLYNQPSDTRQYHENIKINQAMRKKLILYFKRRNHARKQWKQK-FC 299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          64
                                                    -i- DISEASE: Involved in acute myeloid leukemias through a chromosomal translocation t (8;16) (pl1;pl3) involving MYST3 and CREBBP.
-i- SIMILARITY: Contains 2 PHD-type zinc fingers.
-i- SIMILARITY: Belongs to the MYST (SAS/MOZ) family.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | | | ::|| ::|| | ;
QPSRRGRWGRKNRKTQERFGDKDSKLLLEETSSAPQEQYGECGEKSEATQEQYTESEEQL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RPSLLSEFQPGNERSQELHLRPESHSYLPELGKSEMEFIESKRPRLELLPDPLLRPSPLL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REITMVEQQISKLKKKQQQLEEEAAKPPEPEKPVSPPPIESKHRSLVQIIYDENRKKAEA
FUNCTION: May represent a chromatin-associated acetyltransferase.
SUBCELLULAR LOCATION: Nuclear.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BREAKPOINT FOR TRANSLOCATION TO FORM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   translocation; Zinc-finger; Repeat;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EPRYPPHSLSYPVQIARTHTDVGLLEYQHHSRDYASHLSPGSIIQ-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               971 RRYSEGDRAV--LRGFSESSEEE----EEPESPRSSSP-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 2004;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9FFBBAC3792854BA CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2.4%; Score 322.5; DB 1;
larity 19.8%; Pred. No. 0.0013;
Conservative 199; Mismatches 582;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GLN/PRO-RICH,
MET-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PHD-TYPE 1
PHD-TYPE 2
POLY-SER.
C2HC-TYPE.
POLY-GLU.
POLY-ARG.
POLY-ARG.
POLY-LYS.
GLU-RICH.
GLU-RICH.
POLY-RYS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MOZ-CBP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GO; GO:0006333; P:DNA packaging; TA InterPro; IPR005818; Histone H1/H5. InterPro; IPR005717; MOZ SAS. InterPro; IPR001965; Znf PHD. Pfam; PF01853; MOZ SAS; I. Pfam; PF01853; PHD; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; U47742; AAC50662.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SMART; SM00526; H15; 1.
SMART; SM00249; PHD; 2.
PROSITE; PS01359; ZF PHD_1;
PROSITE; PS50016; ZF PHD_2;
Proto-oncogene; Chromosomal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           225054
```

us-09-522-753-5.rsp

	· ·			
300 QRYDQLMEALEKKVERIENNPRRAKESKVREYYEKQFPEIRKQRELQERM 350 1036 ETISTTEVLDEPFEDSDSERPMPRLEPTFEIDEEEEEEDENELFPREYFRRL 1088 351 QSRVGQRGSGLSMSAARSEHEVSEIIDGLSEQENLEKQMRQLAVIPPMLYDADQORIKFI 410	1176 SREIMPVSTQA-CVI	623 AKKGLLEHGRNWSAIARWYGSKTVSQCKNFYFNYKKRONLDEILQOHKIKM-EKERNARR 681 1327 TKKKELEEQPTREDVREEPGVQBSFLDANNGKSREKIKKKEETELD 1372 682 KKKKAPAAASEBAAFPPVVEDEEMRASGVSGNEERWYEBAALHASGNEVPRGECS 737 1373 SEEEQPSHDTSVYSEQMAGSEDDHEEDSHTKEELIELKEEEIFHSELD 1421 738 GPATVNNSSDTESIPSPHTEAAKDTGQNGPKPPATLGADGPPGPPTPPRRTSRAPIEDT 797 1422 LETVQAVQSLTQEESSEHEGAYQDCEETLAACQTLQSYTQADEDPQMSNNVEDC 1474	798 PASEATGAPTPPPAPPSPSAPPPVVPKEEKEEETAAAPPVEEGGEGK 1475 HASEHNSPISSVQSHPSQSVRSVSSPNVPALESGYTQISPEGGSLSAPSMQNMETSPWMD 1534 845 -PPAAEELAVDTGRAEEPVKSECTEEAEGPAKGKDAEAAEATAEGALKAEKKEGGG 901 1535 VPSVSPHSQQVDSGRSDLG-SIESTTENYENPSSYDSTWGGSICGNSSSQSSCSYGGLS 1593 902 RATTAKSGGAPQDSDSSATCSADEVDEAEGGDKWALLSPRPSILTPTGDPRANAS 956 1534 SSSSLTQSSCVVTQQMASMGSSCSMMQQSSVQPAANCSIKSPQSCVVERPPSNQQQ 1650 957PQKPLDLKQLKQRAAAIPPIQVTKVHEPPREDAAPTKPAPPAPPPPQQQPPLSQC 1705	1012 PQQPGSSPRGKSRSPAPPADKEAFAAEAQKLPGDPPCWTSGLPFPVPPREVI- 1063 1706 SMNNSFTPAPMIMEIPESGSTGNISIYERIPGDFGAGSYSQPSATFSLAKLQQLT 1760 1064KASPHAPDPSAFSYAPGHPLPLGLHDTARPVLPRPPTISN 1104 1761 NTIMDPHAMPSSHSPAVTSYSTSVSLSNTGLAQLAPSHPLAGTPQAQATMT 1811 1105 PPPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEHAKAPVQPVTMGLPLPMDPKKL 1161 1812 PPPNLASTTWNLTSPLLQCTMSATNIGIPHTORLQGQM-PVKGHISIRSKSAPL 1864 1162 APPSQVKQPQLSPRGQAGPPESLGVPTAQEASVLRGTALGSVPGGSITKGIPST 1215 1162 APPSGVKQPQLSPRG
6 6 6 6 6 6	8 8 8 8 8	666666	6 6 6 6 6 6 6	8 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6

```
MEDLINE=55207819; PubMed=2581969; Pihlajaniemi T., Tryggvason K., Myers J.C., Kurkinen M., Lebo R., Pihlajaniemi T., Tryggvason K., Myers J.C., Kurkinen M., Lebo R., Cheung M.-C., Prockop D.J., Boyd C.D.; "CDNA clones coding for the pro-alphal(IV) chain of human type IV procollagen reveal an unusual homology of amino acid sequences in two J. Biol. Chem. 260:7681-7687(1985).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE OF 1-943 FROM N.A.
TISSUB-Placenta;
MEDLINE=88029471; PubMed=3311751;
MEDLINE=88029471; PubMed=3311751;
Brazel D., Oberbaeumer I., Dieringer H., Babel W., Glanville R.W.,
"Completion of the anino acid sequence of the alpha 1 chain of human
basement membrane collagen (type IV) reveals 21 non-triplet
interruptions located within the collagenous domain.";
Eur. J. Biochem. 168:529-536 (1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           [4] SEQUENCE OF 28-243.
SEQUENCE OF 28-243.
MEDLINE=86004708; PubMed=4043082;
Glanville R.W., Qian R.Q., Siebold B., Risteli J., Kuehn K.;
"Amino acid sequence of the N-terminal aggregation and cross-linking region (75 domain) of the alpha 1 (IV) chain of human basement
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
MEDLINE=89340433; PubMed=2701944;
Soininen R., Huctari M., Ganguly A., Prockop D.J., Tryggvason K.;
"Structural organization of the gene for the alpha 1 chain of human type IV collagen.";
J. Biol. Chem. 264:13565-13571 (1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE=88083584; PubMed=3691802;
Soininen R., Haka-Risku T., Prockop D.J., Tryggvason K.;
"Complete primary structure of the alpha 1-chain of human basement
membrane (type IV) collagen.";
FEBS Lett. 225:188-194(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE OF 534-1447.

MEDLINE-85003629; PubMed-6434307;

Babel W., Glanville R.W.;

"Structure of human-basement-membrane (type IV) collagen. Complete amino-acid sequence of a 914-residue-long pepsin fragment from the alpha 1(IV) chain.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE OF 1259-1669 FROM N.A.
MEDILINES-85216555; PubMed=2582422;
Brinker J.M. Gudas L.J., Loidl H.R., Wang S.-Y., Rosenbloom J.,
Kefalides N.A., Myers J.C.;
                                    1264 ----PKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGP 1300
                                                                                                                                                                                                                                                                                 21-JUL-1986 (Rel. 01, Created)
01-FEB-1996 (Rel. 33, Last sequence update)
10-OCT-2003 (Rel. 42, Last annotation update)
Collagen alpha 1(IV) chain precursor.
                                                                                                                                                                                                                                   Ą.
                                                                                                                                                                                                                               PRT; 1669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Eur. J. Biochem. 152:213-219(1985)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE OF 46-1257 FROM N.A. TISSUE=Placenta;
                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                               Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         collagen.";
                                                                                                                                                                                                                            CA14 HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         membrane
                                                                                                                                                   RESULT 88

CM14 HUMAN

ID CANA HUMAN

ID CANA HUMAN

AC PO146 AI

DT 01-FEB

DT 10-OCT-

OC BUKATY

OC BUKATY

OC BUKATY

OC BUKATY

OC BUKATY

OC BUKATY

OC BUKATY

OC BUKATY

OC BUKATY

OC BUKATY

OC BUKATY

OC BUKATY

OC BUKATY

OC BUKATY

OC BUKATY

OC BUKATY

OC BUKATY

OC BUKATY

OC BUKATY

OC BUKATY

OC BUKATY

OC BUKATY

OC BUKATY

OC BUKATY

OC BUKATY

OC BUKATY

OC BUKATY

IT SECULE

RA SECULE

RA BEAZE

RA BEAZE

RA BEAZE

RA BEAZE

RA BEAZE

RA BEAZE

RA BEAZE

RA BEAZE

RA BEAZE

RA BEAZE

RA GLANY

RT "Amin

RT "Amin

RT "Amin

RT "Amin

RT BULL

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BABDLI

RA BA
                                                                                                                                                                                                                                                                                                                                                                                                                    COL4A1
                                                  Д
ઠે
```

```
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DISULPID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DISULPID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                  PROPEP
CHAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                         SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                             EMBL;
EMBL;
                                                                                   EMBL;
EMBL;
                                                           EMBL;
EMBL;
                                                                                                           EMBL;
                                                                                                                                             EMBL;
                                                                                                                                                       EMBL;
                                                                                                                                                                                                                   EMBL;
                                    EMBL;
                                                                                                                      EMBL;
                                                                                                                                                                     EMBL;
                                                                                                                                                                                             EMBL;
                                                                                                                                                                                                       EMBL;
   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                         nidgen.

-1-SUBUNIT: There are six type IV collagen isoforms, alpha 1(IV)-
alpha 6(IV), each of which can form a triple helix structure
with 2 other chains to generate type IV collagen network.

-1-DOMAIN: Alpha chains of type IV collagen have a noncollagenous
domain (NCI) at their C-terminus, frequent interruptions of the
G-X-Y repeats in the long central triple-helical domain (which may
cause flexibility in the triple helix), and a short N-terminal
triple-helical 7S domain.

-1-PTM: Lysines at the third position of the tripeptide repeating
unit (G-X-Y) are hydroxylated in all cases and bind carbohydrates.

-1-PTM: Prolines at the third position of the tripeptide repeating
unit (G-X-Y) are hydroxylated in some or all of the chains.

-1-PTM: Type IV collagens contain numerous cysteine residues which
are involved in inter- and intramolecular disulfide bonding. 12 of
these, located in the NCI domain, are conserved in all known type
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the BWBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                              SEQUENCE OF 1-28 FROM N.A.
MEDLINE=89034231; PubMed=3182844;
Solininen R., Huotari M., Hostikka S.L., Prockop D.J., Tryggvason K.;
"The structural genes for alpha 1 and alpha 2 chains of human type IV
collagen are divergently encoded on opposite DNA strands and have an
overlapping promoter region.";
                                                                                                                                                                MEDLINE=8905112; PubMed=2844531; Siebold B., Deutzmann R., Kuehn K.; Siebold B., Deutzmann R., Kuehn K.; Siebold B., Deutzmann R., Kuehn K.; Siebold B., Deutzmann P., Kuehn K.; Siebold B., Deutzmann to fintra- and intermolecular disulfide bonds in the carboxyterminal, non-collagenous aggregation and cross-linking domain of basement-membrane type IV collagen."; Eur. J. Biochem. 176:617-624(1988).

-I- FUNCTION: Type IV Collagen is the major structural component of glomerular basement membranes (GBM), forming a 'chicken-wire' meshwork together with laminins, proteoglycans and entactin/
"Restricted homology between human alpha 1 type IV and other procollagen chains.";
Proc. Natl. Acad. Sci. U.S.A. 82:3649-3653(1985).
                                                                                                                                           SEQUENCE OF 1441-1669, AND DISULFIDE BONDS
                                                                                                                    Biol. Chem. 263:17217-17220(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           JOINED.
JOINED.
JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; M26576; AAA53098.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAA53098.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAA53098.
AAA53098.
AAA53098.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAA53098.
AAA53098.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAA53098
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      IV collagens.
                                                                                                                                                         TISSUE=Placenta;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     M26559;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              M26540;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  M26544;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        J04217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    M26543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         M26546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            126538
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL;
EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MBL;
```

```
100;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      808
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PGIPGCNGTKGERGPLGPPGLPGFAGNPGPPGLPGMKGDPGEILGHVPGMLLKGERGFPG 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       728 GNEVPRGECSGPATVNNSSDTESIPSPHTEAAK-DTGQNG----PKPPATLGADGPPPGP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GMQGPEGP-QGPP--GQKGDTGEPGLPGTKGTRGPPGASGYPGNPGLPGIPGQDG-PPGP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (7S DOMAIN).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 2.4%; Score 320.5; DB 1; Length 1669; Best Local Similarity 21.0%; Pred. No. 0.0013; Matches 415; Conservative 133; Mismatches 750; Indels 677;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SMART; SM00111; C4; 2.
Extracellular matrix; Connective tissue; Basement membrane;
Repeat; Hydroxylation; Glycoprotein; Collagen; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COLLAGEN ALPHA 1 (IV) CHAIN.
TRIPLE-HELICAL REGION.
NONHELICAL REGION (NC1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3BEBA6DFFB9B8A84 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AMINO-TERMINAL PROPEPTIDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED (GLCNAC. . .).
OR 1548.
OR 1551.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      783 PTPP----RRTSRAPIEPTPASEATGAPTP------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         OR 1662.
OR 1665.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ٢
                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam; PF01413; C4; 2.
Pfam; PF01391; Collagen; 24.
ProDom; PD000007; C1g helix; 6.
ProDom; PD003923; ProcollagnC4; 1.
JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
JOINED.
                                                                                                                                                                                                                            JOINED.
JOINED.
                                                                                                                                                                                            JOINED.
                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR008161; Clg helix.
InterPro; IPR008160; Collagen.
InterPro; IPR001442; Procollagn4_
                                                                                                                                                                                JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1358 1358 E -
1669 AA; 160611 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8 0 0 Z
                                                                                               AAA53098.1;
AAA53098.1;
AAA53098.1;
                                                  AAA53098.1;
AAA53098.1;
                                                                                                                                                                                                                                                                                                                             AAA52042.1;
CGHU4B.
   AAA53098.1;
                                                                                                                                                                                                                            AAA53098.1;
                                                                               AAA53098.1;
                                                                                                                                                                                               AAA53098.1;
                                                                                                                                                                                                                AAA53098.1;
                                                                                                                                                                                                                                                              AAA53098.1;
                                                                                                                                                                                                                                                                                                                                          PIR; S16876; CGHU4B.
Genew; HGNC:2202; COL4A1.
                                 AAA53098.1;
                                                                                                                                                                                                                                                                             CAA68698.1
                                                                                                                                                                                                                                                                                              CAA29075.1
                                                                                                                                              AAA53098..
                                                                                                                                                                                                                                                                                                               AAA52006
                                                                                                                                                                                                                                                                                                                                                                            MIM; 120130; -
                                                              M26541;
                                                                                                                                                                                                                                                                                                             M10940;
                                                                                                                                                                                                                                                M26574;
                                                                                                                                                                                                                                                            M26575;
                                                                                                                                                                                                                                                                                             X05561
```

.

OVI------VPPTPGTPATAMD 1756 SSSERERDRDRERDRDREREKSIL 1815 : : | | | | | AGQKGEPGL 1170 SSSRPASHSHAHQHSPISPRTQDA 1870 SPGIPGSKGEQGFMGPPGPQGQPG 1222 STSTSSPVRPAATFPPATHCPLGG 1923 ------PGPMGPPG----LPG 1260 SPDPPAPPASASDPHREKTQSKP 2030 .||| .QGLP-----GPKG 1381 AKPPARSGLEPASSPSKGSEPRP 1983 -PGPPGPPG------- 1351 KGLPKHLEELDKSHLEGELRPK- 2089 KGHQRVVTLAQHISEVITQDYTR 2149 KG----- 1408 PPDHGAPARGSPHSEGGKRSPEP 2209 GPD-----GLPGSMGPPGTPSV 1442 RDG-EQTEPSRMGSKSP--GNTS 2266 | |: VQGNERAHGQDLGTAGSCLRKFS 1498 YNISOPGTEIFNMPAITGTGLMT 2326 ::: | | CPSGWSSLWIGYSFVM----- 1585 SANAFNPLNASAŞLPAAMPITAA 2386 3D-RPPSVSSVHSEGDCN 2440 : | EEFRSAPFIECHGRGTCN 1623 rata; Euteleostomi; asianidae; Phasianinae; ; Q90797; Q92014; gene."; Ä Ann. N.Y. Acad. Sci. 460:85-116(1985). Z.

clone.";

```
"Structure of the promoter for chicken alpha 2 type I collagen gene."; Proc. Natl. Acad. Sci. U.S.A. 78:5334-5338(1981).
                                                                                                                                                                                                                                                                                                                                                                                                                               "Construction and characterization of cDNA clones encoding the 5' end of the chicken pro alpha 1(I) collagen mRNA.";
Gene 56:71-78(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Kang A.H., Igarashi S., Gross J.;
"Characterization of the cyanogen bromide peptides from the alpha-2
chain of chick skin collagen.";
Biochemistry 8:3200-3204(1969).
                                                                                                                                                                                                                                                          SEQUENCE OF 1-33 FROM N.A.
MEDLINE-84297217; PubMed-6473103;
Aho S., Tate V.E., Boedtker H.;
"Location of the 11 bp exon in the chicken pro alpha 2(1) collagen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Highberger J.H., Kang A.H., Gross J.;
Comparative studies on the amino acid sequence of the alpha 2-CB2
peptides from chick and rat skin collagens.";
Biochemistry 10:610-616(1971).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Kang A.H., Gross J.;
"Amino acid sequence of cyanogen bromide peptides from the amino-
terminal region of chick skicollagen.";
Biochemistry 9:796-804(1970).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE OF 566-587 FROM N.A.
MEDLINE=79074829; PubMed=364479;
MEDLINE=79074829; PubMed=364479;
Lebrach H., Frischauf A.-M., Hanahan D., Wozney J., Fuller F.,
Crkvenjakov R., Boedtker H., Doty P.;
"Construction and characterization of a 2.5-kilobase procollagen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-69206882; PubMed-5785233;
Lane J.M., Miller E.J.;
"Isolation and characterization of the peptides derived from the
                SEQUENCE OF 1-89 FROM N.A.
MEDLINE-83246518; PubMed=6135195;
Tate V.E., Finer M.H., Boedkker H., Doty P.;
"Chick pro alpha 2 (1) collagen gene: exon location and coding potential for the prepropeptide.";
Nucleic Acids Res. 11:91-104(1983).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         alpha 2 chain of chick bone collagen after cyanogen bromide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Wozney J., Hanahan D., Tate V.E., Boedtker H., Doty P., "Structure of the pro alpha 2 (I) collagen gene."; Nature 294:129-135(1981).
                                                                                                                                                     MEDLINE=82060240; PubMed=6946474;
Vogeli G., Ohkubo H., Sobel M.E., Yamada Y., Pastan I.,
de Crombrugghe B.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE OF 74-91, 263-448 AND 1088-1169 FROM N.A. MEDLINE=82058081, PubMed=6272119;
                                                                                                                                                                                                                                                                                                                                   gene.";
Nucleic Acids Res. 12:6117-6125(1984)
                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE OF 1-79 FROM N.A.
MEDLINE=88056316; PubMed=3678834;
Finer M.H., Boedtker H., Doty P.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE=71115216; PubMed=5544653;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE=70131186; PubMed=4313735;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE=69285369; PubMed=5809220;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Biochemistry 8:2134-2139(1969).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE OF 78-92 AND 415-448.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE OF 78-92 AND 415-448.
                                                                                                                                        SEQUENCE OF 1-14 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE OF 78-92.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE OF 78-92.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TISSUE=Bone;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               rissue=Skin;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           cleavage."
[2]
SEQUENCE (
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Fuller F., Boedtker H.; "Sequence determination and analysis of the 3' region of chicken proappuence determination and analysis of the 3' region of chicken proalpha 1(1) and pro-alpha 2(1) collagen messenger ribonucleic acids including the carboxy-terminal propeptide sequences."; Biochemistry 20:996-1006(1981).
                                                                                                                                                    MEDLINE=81264246; PubMed=6267043;
Dickson L.A., Ninomiya Y., Bernard M.P., Pesciotta D.M., Parsons J., Green G., Elkenberry E.F., de Crombrugghe B., Vogeli G., Pastan I., Fietzek P.P., Olsen B.R., "The exon/intron structure of the 3'-region of the pro alpha 2(I)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      hydroxyapatite. PTM: Prolines at the third position of the tripeptide repeating unit (G-X-Y) are hydroxylated in some or all of the chains.
                                                                                                                                                                                                                                                                      MEDLINE-81064671; PubMed-6159982;
Avvedimento V.E., Vogeli G., Yamada Y., Maizel J.V. Jr., Pastan I.,
                                                                                                                                                                                                                                                                                                                                                                         "Correlation between splicing sites within an intron and their sequence complementarity with Ul RNA."; Cell 21:689-696(1980).
                                                                                                                                                                                                                                                                                                                                                               SEQUENCE OF 126-161; 467-517 AND 926-954 FROM N.A.
          Proc. Natl. Acad. Sci. U.S.A. 75:5417-5421(1978).
                                                                                                                                            SEQUENCE OF 998-1169 AND 1234-1362 FROM N.A.
                                                                                                                                                                                                                                                         SEQUENCE OF 932-954 AND 968-980 FROM N.A.
                                                                                                                                                                                                                      collagen gene.";
J. Biol. Chem. 256:8407-8415(1981).
                                                   MEDLINE=81160715; PubMed=6927845;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        M25959; AAA69960.1; JOINED.
M25961; AAA69960.1; JOINED.
M25962; AAA69960.1; JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; M25963; AAA69960.1; -.
EMBL; M25956; AAA69960.1; JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           JOINED
                                     SEQUENCE OF 902-1362 FROM N.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAA69962.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             M25964; AAA69961.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               M25965; AAA69961.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAA69962.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAA69962.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAA69962.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAA69962.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAA69962.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAA69962
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAA69962
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAA69962
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAA69962
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAA69962
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAA69962
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAA69962
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAA69962
                                                                                                                                                                                                                                                                                                 de Crombrugghe B.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        M25984;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          M25974;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    M25977
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL;
EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL;
EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL;
EMBL;
EMBL;
EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL;
```

M25979;

OY 1127 QGMSVQLHVPYSEHAKAPVGPV-TMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLG 1185	AIPPERHSPHHIKEQHHIRGSITQGIP 	QY 1406 KLKPAHEGLVATVKEAGRSIHEIPREELRHTPELPLAPRPLKEGSITQGTP-LKYDTG 1462 Db 617 GNKGEPGNVGPAGAPGPGGG-GIPGERGVAGVPGGKGEKGAPGLRGDTG 665 QY 1463 ASTTGSKKHDVRSLIGSPGRTFPPVHPLDVMADARALERACYEESLKSRPGTASSGG 1519 Db 666 ATGRDGARGLPGAIGAPGPAGGAGDRGEGGPAGPAG 701	QY 1520 GSIARGAPVIVPELGKBRQ-SPLTYEDHGAPFAGHLPRGSPVTMREPTPRLQEGSLSSSK 1578 b 10 1 1 1 1 1 1 1 1 1	QY 1639 RGIP-LDAAAAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAALENRQTIINDYITSQ 1693	OY 1754 AMDRIAYLPTAPOPFSSRHSSPLSPGGPTHLTKPTTTSSSERERDRDRERDRDREREKS 1813	Qy 1871 LQQRPSVLHNTGMKGIITAVEPS-KPTVLRSTSTSSPVRPAATFPPATHCPLGGTLDGVY 1929 Db 958 LGAPGPHGQVGPSGRFGNRGDPGPVGPAGAFGPR	Qy 1990 GHATIARTPAKNIAPHHASPDPPAPPASASDPHREKTQSKPFSIQELELRS 2040 Db 1022 GHNGLQGLPGLAGQHGDQGPPGNNGPAGPRGPPGRDGRDGRDGRDGPGGPRGPRGPRGPRGPRGPRGPRGPRGPRGPRGPRGPRG	OY 2097 GGEAAHLPHLRPLPESQPSSSPLLQTAPGWKGHQRVVTLAQHISEVITQDYTR 2149
DR EMBL; M25980; AAA69962.1; JOINED. DR EMBL; M25981; AAA69962.1; JOINED. DR EMBL; M25981; AAA69962.1; JOINED. DR EMBL; M25983; AAA69962.1; JOINED. DR EMBL; M000826; AAA451611.1; JOINED. DR EMBL; J000826; AAA451611.1; JOINED. DR EMBL; J000810; AAA51611.1; JOINED. DR EMBL; J000830; AAA51611.1; JOINED. DR EMBL; J000830; AAA51613.1; JOINED. DR EMBL; J000830; AAA51613.1; JOINED. DR EMBL; J000812; AAA51613.1; JOINED. DR EMBL; J000812; AAA51613.1; JOINED. DR EMBL; J000812; AAA51613.1; JOINED. DR EMBL; J000813; AAA51613.1; JOINED. DR EMBL; J000811; AAA51615.1; JOINED.	J00814, AAAS1615.1, J00815, AAAS1615.1, K02657, CAA26493.1, K00794, -; NOT ANNOT V00390; CAA2368.1, MITC88, AAA48673.1, MIC81, AAA48633.1, MIC81, AAA48633.1, MIC81, AAA48633.1,	EMBL; J00828; AAA516 EMBL; J00821; -, NOT EMBL; J00832; -, NOT EMBL; J00833; -, NOT EMBL; J00833; -, NOT EMBL; J00832; -, NOT PIR; ISO173; ISO173; PIR; ISO206; CGCH2S.	DR INTERPROPT IPROBAGA; Calgapen. DR INTERPROPTO; IPROBAGGS; Callagen. DR PRODOM; PD000007; Clg_helix; 4. DR PRODOM; PD003078; Fib_collagen_C; 1. SMCDOM; PD0038; COLFF; 1. SMCAT; SMCOD38; COLFF; 1. KW SKRATCALIULAR MARTIX; Connective tissue; Repeat; Hydroxylation; KW Glycoprotein; Collagen; Signal.	20 % G-15			AGLPG VIKAS	KHPSVLERQIGAIS

8 6 8 6 8 6 8 6 8 6 8 6 8 6

```
Best Local Similarity 20.8%; Pred. No. 0.0015;
Matches 405; Conservative 153; Mismatches 739; Indels 652; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          "Cell Biology." (Cell Biology.") Agris 73-1582 (1996).

-!- FUNCTION: May play an important role in fibrillogenesis by controlling lateral growth of collagen II fibrils (By similarity).

-!- SUBUNIT: Trimers composed of three different chains: alpha 1(XI), alpha 2(XI), and alpha 3(XI). Alpha 3(XI) is a post-translational modification of alpha 1(II). Alpha 1(V) can also be found instead of alpha 3(XI)=1(II) (By similarity).
                                                                                                                                                                                                                                                                                                                                                             Vandenberg P., Vuoristo M.M., Ala-Kokko L., Prockop D.J.;
"The mouse collia2 gene. Some transcripts from the adjacent rxr-beta gene extend into the collia2 gene.";
Matrix Biol. 15:359-367(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              rumaki N., Kimura T., Matsui Y., Ochi T.;
"Separable cis-regulatory elements that contribute to tissue- and
site-specific alpha 2(XI) collagen gene expression in the embryonic
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Tsumaki N., Kimura T.;
"Differential expression of an acidic domain in the amino-terminal propeptide of mouse pro-alpha2(XI) collagen by complex alternative splicing ";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Name=6; Synonyms=E59;
IsoId=064739-7; Sequence=VSP_007345, VSP_007347;
                                                                                                                                                                      Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Butheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
                                                                                                                                                                                                                       SEQUENCE FROM N.A. (ISOFORM 7).
STRAIN=129/SvJ;
ROWEN L., Qin S., Madan A., Loretz C., James R., Dors M., Mix L.,
Hall J., Lasky S., Hood L.;
"Sequence of the mouse major histocompatibility locus class II
region.";
Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE OF 1-624 FROM N.A. (ISOFORMS 1; 2; 3; 4; 5 AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Name=5; Synonyms=E589;
IsoId=064739-6; Sequence=VSP_007345, VSP_007346;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IsoId=Q64739-5; Sequence=VSP_007346, VSP_007347;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Event=Alternative splicing; Named isoforms=7;
Comment=Additional isoforms seem to exist;
                                                                            0647<u>7</u>9; 061432; 0921W0; FRI; 1/3b A4.
15-JUL-1998 (Rel. 36, Created)
10-OCT-2003 (Rel. 42, Last sequence update)
10-OCT-2003 (Rel. 42, Last annotation update)
Collagen alpha 2(XI) chain precursor.
                                                                        PRT; 1736 AA
                                                                                                                                                                                                                                                                                                                          SEQUENCE OF 1-1678 FROM N.A. (ISOFORM 7).
STRAIN=FVB/N, and 129/Sv; TISSUE=Cartilage;
MEDLINE=97135795; PubMed=8981332;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 IsoId=064739-2; Sequence=Displayed;
Name=2; Synonyms=E5689;
IsoId=064739-3; Sequence=VSP_007346;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    IsoId=Q64739-4; Sequence=VSP_007345;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN=129/Sv; TISSUE=Liver;
MEDLINE=96427460; PubMed=8830784;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Biol. Chem. 270:2372-2378(1995)
                                                                                                                                                                                                                                                                                                                                                                                                                                        STRAIN=129/Sv;
MEDLINE=95138212; PubMed=7836472;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Name=1; Synonyms=E56789;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Name=3; Synonyms=E5789;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Name=4; Synonyms=E569;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE OF 1-8 FROM N.A.
                                                                        STANDARD;
                                                                                                                                                             Mus musculus (Mouse)
                                                                                                                                                                                              NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  mouse cartilage.
           1151 KNPAR-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Name=7;
                                                                        CA2B MOUSE
           윱
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED (GLCNAC. .) (POTENTIAL).
Missing (in isoform 3, isoform 5, isoform
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Missing (In isoform 2, isoform 4, isoform 5 and isoform 6).
IsoId=064739-1; Sequence=VSP 007345, VSP_007347;
-!- PTM: Prolines at the third position of the tripeptide repeating unit (G-X-Y) are hydroxylated in some or all of the chains.
-!- SIMILARITY: BELONGS TO THE TERRILLAR CLASS OF COLLAGENS.
-!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Missing (In isoform 4, isoform 6 and isoform 7).

/FIIdavSP 007347.

R -> L (IN REF 2).
P -> S (IN REF 2).
V -> A (IN REF 2).
TGP -> HGS (IN REF 2).
TGP -> HGS (IN REF 2).
TGP -> HGS (IN REF 2).
TGP -> HGS (IN REF 2).
TGP -> HGS (IN REF 2).
TGP -> HGS (IN REF 2).
TGP -> HGS (IN REF 2).
TGP -> HGS (IN REF 2).
TGP -> HGS (IN REF 2).
TGP -> HGS (IN REF 2).
TGP -> HGS (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
TGP -> U (IN REF 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          matrix; Connective tissue; Repeat; Hydroxylation; Collagen; Alternative splicing; Signal.
1 22 POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COLLAGEN ALPHA 2(XI) CHAIN.
CARBOXYL-TERMINAL PROPEPTIDE.
TSP N-TERMINAL.
NONHELICAL REGION.
TRIPLE-HELICAL REGION.
NONHELICAL REGION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              18D792D4A3387C61 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /FTId=VSP_007345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                6 and isoform
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; U16789; AAA67751.1;
EMBL; U16789; AAA67752.1;
EMBL; D34042; BAA18310.1;
EMBL; D34066; BAA12208.1;
EMGD; MGI:88447; CO11142.
InterPro; IPRO08161; Clg helix.
InterPro; IPRO08160; Collagen.
InterPro; IPRO0885; Fib Collagen.
InterPro; IPRO0885; Fib Collagen.
InterPro; IPRO0885; Fib Collagen.
InterPro; IPRO01791; Laminin.G.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam; PF01410; COLFI; 1.
Pfam; PF01391; Collagen; 18.
Pfam; PF02210; TSPN; 1.
ProDom; PD000007; Clg helix; 1.
ProDom; PD002078; Fib collagen_C; 1.
SMART; SM00038; COLFI; 1.
SMART; SW00282; LamG; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     'FTIG=VSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Extracellular matrix; Connective
                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; AF100956; AAC69905.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR003129; TSPN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 536
621
705
797
845
854
876
889
922
1005
11386
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           313
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3lycoprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
```

101; 651

DB 1; Length 1736;

2.4%; Score 319;

600 QSAELASMELNESSRWTEEEM-----ETAKKGLLEHGRNWSAIARWVGSKTVSQCKN

Db 1004NEGPAGPPGPAGSPGERGAAGSGGPIGPPGRPGPQGPPG 1042	1561 TWREPTPRLQEGSLSSSKASQDRKLTSTPREIAKSPHSTVPEHHPHPISPYEHLLRGVSG 1620	1621 VDLYRSHIPLAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAALE 1680	1681 NRQTIINDYITSQQMHHNTATAMAQRADMLRGLSPRESSLALNYAAGPRGIIDLSQVPHL 174	1741 PVLVPPTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTHLTKPTTTSSS	1174	Y 1795 ERERDRDRERDRDREREKSILISTITVEHAPIWRPGTEQSSGSSGGGGGSSSRPA 1852 : :	Y 1853 SHSHAHQHSPISPRTQDALQQRPSVLHWTGMKGIITAVEPSKPTVLRSTSSPVRPAAT 1912	1913 FPPA-THCPLGGTLDGVYPTLMEPVLLPKEAPRVARPERPRADTGHAFLAKPPARSGLE	1971 -PASSPSKGSEPRPLVPPVSGHATIARTPAKNLAP	1370	7 2005 HHASPDPPAPASASDPHREKTÖSKPFSIQELELRSLGYHGSSYSPEGVEPVSPVSSPSL 2064	2065 THDKGLPKHLEBLDKSHLEGELRPKQPGPVKLGGEAAHLPHLRPLPESQP 211	2115SSSPLLQTAPQVKGHQRVVTLAQHEVI-TQDYT	1513	2162 PLYSFPGASCPVLDLRRPPSDLYLPPPDHG	2192APARGSPHSEGGKRSPEPNKTSVLGGGEDGIEPVSP 2227	2228 PEGWIEPGHSRSAVYPLLYRDGEQTEPSR 2256	ULT 91 CHIRB	GPI CHLAKE STANDARD; PRT; Q9FPQ6; Q03927; 28-PEB-2003 (Rel. 41, Created)		glycoprotein 1). Gpl. Chlamvdomonas reinhardtii	Eukaryota, Viridiplantae, Chlorophyta, Chlorophyceae, Volvocales, Chlamydomonadaceae; Chlamydomonae. NCRT Tarin-10ff.
<u> </u>	<i>&</i> 8	& A 	<i>ે</i> જે	δ	ପ୍ର (`	& ₽ -	<i>ර</i> සි	à	ପ୍ର	<i>장</i> 옵	<i>장</i> 옵	<i>&</i>	<u>ස</u>	상 옵	<i>></i> 옵	S da	RES GP1	PAE		S G B	8888
: : :: : 107 QGVQQLGLELGRPVRELYEDQRGRPQASAQPIFRGLSLADGKWHHVAVAVKGQSVT 162	652 FYFNYKKRQNLDBILQQHKLKMEKERNARRKKKKAPAAASEBAAFPPVVEDEEM 705 :: : :: : 1 1 1 1	CSGPATVANSSDTESIPSPHTBAA	PSPSAP		298EEEGGVLESSPLPFLEEEGTDLQVSPTADSFQAEEYGEGGTDSPAGFYDYTYGYGD 352	865 ECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSSGA 911 :		DLKOLKQRAAAIPPIQVTKVHEPPREDAAPTKPAPPAPPPPQNLQPESDAPQQPGSSP	1020 RGKSRSPAPPADKEAFRARAQKLPGDPPCF 1055	522 KGPQGLTGPPGKAGRRGRAGADGARGMPGEPGMKGDRGFDGLPGLPGEKGQRGDTGAQGL 581	1056 PVPPREVIKASPHAPDPSAFSYAPPGHPLPLGLHDTARPVLPR-PPTISNPPPLISSAKH 1114 1	LPMDPKKLAPFSGVKQEQLSP	1175 KGQAGPPESLGVPTAQEASVLKGTALGSVPGGSITKGIPS 1214	642 QGEPGPPGQQGTPGAQGLPGPQG-AIGPHGEKGARGKPGLPGMPGSDGLPGHPGKEGPPG 700	1215 TRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPSRLDRGREDSLPKGHVI 1269 			TEAYK		907 GQRGEVGFQGKTGPPGPPGVVGPQGTAGESGPMGERGHSGPPGPPGECGLPGT 959	960 SGREGTRODPGPPGAPGKDGPAGLRGFPGERGLPGTAGGPGLKG- 1003	
qa	ò 8	<i>&</i> ₽	රු සි	ò	셤	δ O	Ş 5	8 6 8	ò	QQ	è a	8 8	ò	QQ	ර් සි	강 옵	දු දු	oy G	ò	ብ ሪ	7 A	ò

eae; Volvocales; xyproline-rich OX NCBI_TaxID=3055;

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation. the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to licensee@isb.sib.ch).
323 PSPAPPSPAP---SPPPSPAP----PTPSPSPSPSPSPSPSPSPSPSPSPSPSPSPIP---- 371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DECOLINE FROM N.A. Martin E.C., Sharp E., Adler P.N.;

Brunk B.P., Martin E.C., Sharp E., Adler P.N.;

Brunk B.P., Martin E.C., Sharp E., Adler P.N.;

Brunk B.P., Martin E.C., Sharp E., Adler P.N.;

"Drosophila genes Posterior Sex Combs and Suppressor two of zeste encode proteins with homology to the murine bmi-1 oncogene.";

Nature 353:351-353(1991).

-I. FUNCTION: The Polycomb group (Pc-G) genes are needed to maintain expression patterns of the homeotic selector genes of the antennapedia (Antp.C) and bithorax (Bx-C) complexes, and hence for the maintenance of segmental determination.

-I. SUBCELLULAR LOCATION: Nuclear (Probable).

-I. SIMILARITY: Contains 1 RING-type zinc finger.
                                                                                                                                                                                                                                                                                                   1124 AISQGMSVQLHVPYSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPES
                                                                                                     ------SPSPKPSPSPVAVKLVWADDAIAFDDLNGTSTRPGSASRMVGEPDI
                                                                                                                                                                                                              418 AGTKCKGNLKGWMPKPSRNPRWGQAVFSGGRTVGSVANVTIRVAFATEKPALIYSSIELV
                                                                                                                                                                                                                                                                     --TKGIPSTRVP----SDSAITYRGSITHGTPADVLYKGTIT-----RIIGEDSPSR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       :ein; DNA-binding; Nuclear protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Drosophila melanogaster (Fruit fly).
Bukaryota, Metazoa, Arthropoda, Hexapoda; Insecta; Pterygota;
Neoptera; Endopterayota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
                                                                                                                                                            1184 LGV-------PTAQEASVLRGTALGSVPGGSI-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-JUN-1994 (Rel. 29, Created)
01-JUN-1994 (Rel. 29, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT; 1603 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POLY-THR.
POLY-THR.
POLY-THR.
POLY-THR.
POLY-THR.
POLY-SER.
POLY-SER.
POLY-SER.
POLY-SER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PROSITE, PS00518; ZF RING 1; 1. PROSITE, PS50089; ZF RING 2; 1. Zinc-finger; Developmental prot
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FlyBase; FBgn0005624; Psc.
InterPro; IPR001841; Znf_ring.
Pfam; PF00097; zf-C3HC4; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3MBL; X59275; CAA41965.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Posterior sex combs protein.
                                                                                                                                                                                                                                                                                                                                                                        1255 LDRGREDSLPK 1265
                                                                                                                                                                                                                                                                                                                                                                                                         534 INMGAGNKKPK 544
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 S17983; S17983
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=7227;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Zinc-finger;
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DROME
                                                                                                                                                                                                                                                                     1209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               OMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 92
PSC_DROME
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8
                                                                                                     용
                                                                                                                                                 ò
                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                     ઠ
                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                                                                                     8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       946 TPTGDPRANASPQKPLDLKQLKQRAAAIPPIQVTKVHEPPREDAAPTKPAPPAPPPPPQNL 1005
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1006 QPESDAPQQPGSSPRGKSRSPAPPADKEAFAAEAQKLPGDPPCWTSGLPFPVPPREV-IK 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1065 ASPHAPDPSAFSYAPPGHPLPLGLHDTARP-VLPRPPTISNPPPLISSAKHPSVLERQIG 1123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          717 MVEEAEALHASGNEVPRGECSGPATVNNSSDTESIPSPHTEAAKDTGONGPKPPATLGAD 776
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        233 PPSPSPAPPSPVPP-----SPAPPSPAPSPKPP---APPPPPSPPPPPRPP 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GPPPGPPTP----PRRTSRAPIEPTPASEATGAPTPP-PAPPSPSAP-----PPVVPKE 825
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EKEEETAAAPPVEEGEEQKPPAAEELAVDTGKAEEPVKSECTEEAEEGPAKGKDAEAAEA 885
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SPPSPAPPLPPSPAPPSPSPVPP----SPSPPVPPSPAPPSPTPPSPSPA 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              886 TAEGALKAEKKEGGSGRATTAKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSLL 945
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WVVLAFVASANAQCVPGGIFNCPPSPAPPSPAPPSPAPPSPAPPSPAPPSPGPPSPAPPS 77
                                                                                                                                    ij
                                                                                                                                                                                                                                                                                                                                          encoding cell wall hydroxyproline-rich glycoproteins.";
Proc. Natl. Acad. Sci. U.S.A. 87:7355-7359(1990).
-!- FUNCTION: Major component of the outer cell wall W6 (crystalline)
                                                                                                                         "Glycosylated polyproline II rods-with-kinks as a structural motif plant hydroxyproline-rich glycoproteins."; Biochemistry 40:2978-2987(2001).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VEGETATIVE CELL WALL PROTEIN GP1.
49 X 5 AA APPROXIMATE PPSPX REPEATS.
                                                                       Ferris P.J., Woessner J.P., Waffenschmidt S., Kilz S., Drees J.,
Goodenough U.W.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED (GLCNAC. .) (POTENTIAL)
N-LINKED (GLCNAC. .) (POTENTIAL)
N-LINKED (GLCNAC. .) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2.4%; Score 317; DB 1; Length 555;
24.4%; Pred. No. 0.00059;
Live 39; Mismatches 277; Indels 146;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ..
                                                                                                                                                                                                                                                             MEDLINE=91017504; PubMed=1699225;
Adair W.S., Apt K.E.;
"Cell wall regeneration in Chlamydomonas: accumulation of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             93 N-LINKED (GLCNAC. . .) (PC
54219 MW; 6A584A90465502F5 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                  SUBUNIT: Associates with GP2 and GP3. PTM: N-glycosylated and O-glycosylated.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL
                                                                                                                                                                                                                                       PARTIAL PRELIMINARY SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; AF309494; AAG45420.1; -.
EMBL; M58496; AAA69706.1; ALT_SEQ.
G]yccouiteDB; Q9FPQ6; -.
InterPro; IPR003882; Pistil extensin.
PRINTS; PR01218; PSTLEXTENSIN.
                           SEQUENCE FROM N.A.
MEDLINE=21159092; PubMed=11258910;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Glycoprotein, Repeat, Signal SIGNAL 1 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      555
333
273
455
455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      30 55
40 3
259 2
399 3
455 4
493 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity
Matches 149; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                  laver
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
DOMAIN
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   777
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        926
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY RRYRY R
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   음
```

ద ઠ å ð

ò

ઠે 요 ઠે

us-09-522-753-5.rsp

1016	OY . 1060 KEVIKASPHAPDPSAFSYAPP 	Qy 1120 RQIGAISQGMSVQLHVPYSEH :: Db 1219SPRVITPPKTS:	Qy 1180 PPESLGVPTAQEASVLRGTALG 	1240	Db 1302 RPILPŘ	1316	Qy 1354 SITQGIPRSYVEAQED)	Qy 1409 PAHEGLVATVR	Db 1408 PWHSQNAPLSIASSASAAAV	Qy 1456 PLKYDTGASTTGSKKHDVRSLI : : Db 1465 NMWYN	Qy 1514TASSSGGSIA	Db 1500 NEVFNDYLQKLKTAAATGG	Qy 1562 MREPTPRLQEGSLS : : : : : : Db 1550 SPLPAGKLTAAATAPQTKGNSS	RESULT 93	12	AC P28481; DT 01-DEC-1992 (Rel. 24, Created) DT 01-DEC-1992 (Rel. 24, Last sequent	Collagen alpha 1 (Col.2A)		Mammaila; butnerla; Rodentla; NCBI_TaxID=10090; [1]	SEQUENCE FROM MEDLINE=913584	Metsaran "Mouse ty structur	RL J. Biol. Chem. 266:16862-16869(:		RT "Specific hybridization probes of collagen mRNAs.";	RL Biochim. Biophys. Acta 1089:241.
1214 1217 1391 1396 1458 1461 1517 1520 1603 AA; 169999	Query Match 2.4%; Score 316; DB 1; Length 1603; Best Local Similarity 18.4%; Pred. No. 0.0017; Matches 296; Conservative 192; Mismatches 524; Indels 596; Gaps 70;	292 NHARKQWKQKFCQRYDQLMEALEKKVERIENNPRRAKESKVREYYEKQPPEIRKQRELQ 347	348 ERMOSRVGQRGSGLSMSAARSEHEVSEIIDGLSEQENLEKQMRQLAVIPPMLYD 401	ADQQRIKFINMNGLMADPMKV-		450 NFGLIASFDERKTYARCYLYYYLTKKNBNYKSLYRRSYRRRGKSQQQQQQQQQQQQQ 507 				ш-	LISSASSISKISEUTI-VVSSPISEPNIK-LKIDLSKQNSVTIIDMSDPERREIVKPLKPEKE RNARRKKKKAPAASFEAAEDDINGNDS	SRSKKKDKDGSPKSSSSSSSSSSGERKRKSPSPL	719 EBAEALHASGNEVPRGECSGPATVNNSSDTESIPSPHTEAA-KDTG 763	ONGP	743 NNLDDSILMKPPSCMPPKSIASSKRKSKBPVKAVSKKQKLSPPLPTVDFKIRLPV 797	798 PASEATGAPTPPPAPPSPSAPPPVVPKEEKESETAAAPPVEEGEEQKP 845 ::	, α		892 KAEKKEGGSGRATTAKSSGAPQDSDSSATCSADE 925		YNIPTWPTYK	HEPPREDAA :	YTPKPTPNSGSGN	1005 1015 1067 GGGSSSATOSGGNNINTVININTVININTVINIDADE SEPTEMBER 1000 1000 1000 1000 1000 1000 1000 10	
FT FT FT SQ	У М 2.	S G	& 8	ò	g .	<u>8</u>	\$ E	8 8	셤	ठ र	3 &	q	8 6	ò	đ	୪ ଶ	ò	q _Q	ð 8	8	qq	δ	යි දි	Š. 8	

```
----RSPAPPADKEAFAAEAQKLPGDPPCWTSGLPFPVPP 1059
                     IEIVKLPDQPQDQVQAAKEAQKRQSPPAAVPGHLAAKLPP 1186
                                                    PGHPLPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLE 1119
                                                                     HAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAG 1179
                                                                                                                               LGSVPGGSITKGIPSTRVPSDSALTYRGSITHGTPADVLY 1239
                                                                                                                                                                                            EDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSG 1299
                                                                                                                                                                                                                                                                          RAISSASIEGLMGRAIPPERHSP-----HHLKEQHHIRG 1353
                                                                                                                                                                                                                                                                                                                                                                                   KEAGRSIHEIPREELR--HTPELPLAPRPLKEGSITQGT 1455
                                                                                                                                                                                                                                                                                                                                                                                                      IGSPGRIFPPVHPLDVMADARALERACYEESLK--SRPG 1513
                                                                                                                                                                                                                                                                                                                                                                                                                                                           IARGAPVIVPELGKPROSPLTYEDHGAPFAGHLPRGSPVT 1561
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Craniata, Vertebrata; Buteleostomi,
Sciurognathi; Muridae, Murinae; Mus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               s13;
combrugghe B., Vuorio E.;
clippete nucleotide sequence, exon
icing.";
(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     184;
rombrugghe B., Vuorio E.;
s for mouse type I, II, III and IX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         guence update)
lotation update)
scursor [Contains: Chondrocalcin].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SSKASODRKLTSTPREIAKSPHSTVP 1601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ATIVE SPLICING.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT; 1459 AA
```

```
976 IQVTKVHEPPREDAAPTKPAPPAPPPPQNLQPESDAPQQPGSSPRGKSRSPAPPADKEAF 1035
                                                                                                                                                              1036 AAEAQKLPGDPPCWTSGLPFPVPPREVIKASPHAPDPSAFSYAPPGHPLPLGLHDTARPV 1095
                                                                                                                                                                                                                                                                                                                                             MDPKKLA---PFSGVKQEQLSPRGQAGPPESLGVPTAQEASVLRGTALGSVPGGSITKGI 1212
                                                                                                                                                                                                                                                                                                                                                                                                                                   PSTR-VPSDSAITYRGSIT--HGTPADVLYKGTITRIIGEDSPSRL------DRGR-- 1259
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1309 RTYDMMEGRVGRA----ISSASIEGLMGR-----AIPPERHSPHHLKEQHHIRG-SIT 1356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1414 LVATVKEAGRSIHEIPREE-----LRHTPELPLAPRPLKEGSI-----TQGTP 1456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ----LKYDTGASTTGSKKHDVRSLIGSPGRTFPP------VHPLDVMADARALERACY 1504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RSHIPLAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAALENRQT 1684
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QRGERGFPGLPGPSGEPGKQGAPGASGDR-------GPPGPVGPPGLT-- 1000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1505 BESLKSRPGTASSSGGSTARGAPVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSPVTMRE 1564
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PTPRLOEGSLSSSKASQDRKLTSTPREIAKSPHSTVPEHHPHPISPYEHLLRGVSGVDLY 1624
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ------GVKG------DRGETGALGA 1038
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  --GTLDGVYPTLMEPVLLPKEAPRVARPERPRAD--- 1954
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PGERGAPGNRGFPGQDGLAGPKGAP-----GERGPSGLAGPKGANGDPGRPG 524
    1685 IINDY----ITSQQMHHNTATAMAQRA----DMLRGLSPRESSLALNYAAGPRGIIDL--
                                                                                                                                                                                                          ---- GNPGSPGPAGASGNPGTDGIPGAKGSAGAPG
                                                                                                                                                                                                                                                    LPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEHAKAPVGPVTMGLPLP
                                                                                                                                                                                                                                                                                                                                                                                       LGPKGQAGEPGIAGFKGDQ-GPKGETGPAGPQGAPGPAGEEGKRG-ARGE-PGGAGPIGP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   525 EPGLPGARGL-TGRPGDA-GPOGKVGPSGAPGEDGR--PGPPGPOGARGOPGVMGFPGPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GA----NGEPGKAGEKGLAGAPGLRGLPGKDGETGAAGPPGPSGPAGERGEQGAPGPSGF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QGLP------GPPGPPGEGGKQGDQGIPGEAGAP-----G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LVGPRGERG-----FPGERGSPGAQGLQGPRGLPGTDGP--KGAAGPDGPPGAQGPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GLQGMPGERGAAGIAGPKGD-RGDVGEKGPEGAPGKDGGRGLTGPIGPPGPAGA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -NGEKGEAGPPGPSGSTGARGAPGEPGETGPP--GP-----AGFAG--PPGAD---GQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PGAKGDQG-----EAGQ-----KGDAGAPGPQGPSGAPGPQGP-----TGVTGPKGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RG----AQGP---PGATGFPGAAGRVGPPGANGNPG------PAGPPGPAGKDGPKG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VRGDSGPPGRAGDPGLEGPAGAPGEKGEPGDDGPSGLDGPPGPQGL---AGQRGIVGLPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ----SQVPHL--PVLVPPTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTHLTKP
                                                                                                                                                                                                                                                                                                -----RGPPGP--OGATGP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        QGIPRSYVEAQEDYLRREAKLLKREGTPPPPPP----SRDLTEAYKTQALGPLKLKPAHEG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TITSSSERERDRDRERDRDREREKSILISTITVEHAPIWRPGTEQSSGSSGSGGGGGSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ------GPAGEPGREGSPGADGPPGRDGAA-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SRPASHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGIITAVEPSKPTVLRSTSSPVR
                                                                                                                  AGAAGAR---GNDGQPGPAGPPGPVGPAGGPGFPGAPGAKGEAGPTGARGP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETAA---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1909 PAATFPPATHCPLG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        - EGAQGSRGEP
                                                                                                                                                                                                                                                                                                IAGAPGFPGP-
                                                                                                                                                                                                                                                    1096 ]
                                                                                                                                                                                                                                                                                                                                             1156
                                                                                                                                                                                                                                                                                                                                                                                       421
                                                                                                                                                                                                                                                                                                                                                                                                                                   1213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1565
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1625
                           566
                                                                                                                                                                                                                                                                                                399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              581
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    637
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 719
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  817
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          859
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                903
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         960
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1849
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1025
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                             요
                                                                                                                                                         ò
                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                   ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          원. 장
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its most by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               79;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         765
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APGPRGRDGEPGTPGNPGPAGPPGPPGLSAGNFAAQMAGGYDEKAGGAQMGVMQGPM 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          - PTPPPAPPSPS-APPVVPKEEKEETAAAPPVEEGE-----EQKPPAAEELAVDT 855
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DSSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRANASPQKPLDLKQLKQRAAAIPP 975
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SGN---EEEMVEEAEALHASGNEVPRGECSG--PATVNNSSDTESIPSPHTEAAKDTGQN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TGNVLCDDIICEDPDCLNP---EIPPGECCPICPADLATASGRKLGPKGQKGEPGDIRDG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GKABEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSSGAPQDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GPKP-----PATLGADGP--PPGPPTRTS---RAPIEPTPASEATGA-----
FUNCTION: Collagen type II is specific for cartilaginous tissues.
SUBUNIT: Homotrimers of alpha 1(II) chains.
ALTERNATIVE PRODUCTS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                     repeating
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        tissue; Repeat; Hydroxylation;
                                                                                                                                                     Incompage 278481-2; Sequence-VSP 001139, VSP 001140; PTM: Prolines at the third position of the tripeptide repeaunit (G-X-Y) are hydroxylated in some or all of the chains
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches 353; Conservative 114; Mismatches 620; Indels 575;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 1; Length 1459;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GPMGPRGPPGPAGAPGP-----QGPQGNPGEPGEPGVSGPMGPRGPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TRIPLE-HELICAL REGION.
NONHELICAL REGION (C-TERMINAL)
0 -> R (in isoform Short).
/FILG-WSP 001139.
Missing (in isoform Short).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    F6C84FA7C532E7F2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (BY SIMILARITY).
COLLAGEN ALPHA 1(II) CHAIN
CHONDROCALCIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AMINO-TERMINAL PROPEPTIDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Glycoprotein; Collagen; Signal; Alternative splicing. SIGNAL 1 25
                                                                     Event=Alternative splicing; Named isoforms=2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 315.5; DB
Pred. No. 0.0016;
                                                                                        Name=Long;
IsoId=P28481-1; Sequence=Displayed;
                                                                                                                                                                                                                             SIMILARITY: Contains 1 VWFC domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam; PF01410; COLF1; 1.
Pfam; PF01391; Collagen; 18.
PfoDom; PF0000007; Clg_helix; 5.
ProDom; PD000007; Clg_helix; 5.
ProDom; P0002078; Pib_collagen_C; 1.
SMART; SM00214; VWC; 1.
PROSITE; PS01208; VWFC_1; 1.
PROSITE; PS501208; VWFC_2; 1.
EXtracellular matrix; Connective tissue
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FTId=VSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; X57982; CAA41047.1; -. MGD; MGI:88452; COl2al.
InterPro; IPR008161; Clg helix.
InterPro; IPR008160; Collagen.
InterPro; IPR000885; Fib_collagen_C.
InterPro; IPR00100; VWF_C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   21.28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; M65161; AAA68100.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1459 AA; 139154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GPAGKP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         711
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       916
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        908
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                856
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PROPEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PROPEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
```

718

1734

셤

ò

8 6 8 g

ઠ 셤 ò

```
8
               1039 PGAPGPPGSPGPAGPTGKQGDRGEAGAQGPMGPS-----GPAGARGIAGPQGPRGDKGE 1092
                                        -----TGHAFLAKPPARSGLEPASSPSKGSEPRPLVPPVSGHATIARTPA 1999
                                                                                                                                                                                                            2060 SSPSLTHDKGLPKHLEELDKSHLEGELRPKOPGPVKLGGEAAHLPHLRPLPESQPSSSPL 2119
                                                                                                                                                                    1136 -----GKDGSNGIPGPIGPPGPR 1162
                                                                                                                                                                                                                                                                                            2120 LQTAPGVKGHQRVVTLAQHISEVITQDYTRH--HPQQLSAPLPAPLYSFPGASCPVLDLR 2177
                                                                                                                                                                                                                                                                                                                                    QY----MRADEADSTLRQHDVEV---DATLKSLNNQIBSIRSPDGSRKNPARTCQDLKLC 1261
                                                                                                                          2000 KNLAPHHASPDPPAPPASASDPHREKTQSKPFSIQELELRSLGYHGSSYSPEGVEPVSPV
                                                                                                                                                                                                                                           1163 GRSGETGPVGPPG------SPGPPGPPGPPGPGTDMSAFAGLGQREKGPDPM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE OF 1-463 FROM N.A.
MEDILINE-89123368; PubMed=2914927;
Woodbury D., Benson-Chanda V., Ramirez F.;
Wanino-terminal propeptide of human pro-alpha 2(V) collagen conforms to the structural criteria of a fibrillar procollagen molecule.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE OF 398-1496 FROM N.A.
MEDLINE=87146331; PubMed=3029669;
Weil D., Bernard M.P., Gargano S., Ramirez F.;
"The pro alpha 2(V) collagen gene is evolutionarily related to the major fibrillar-forming collagens.";
Nucleic Acids Res. 15:181-198(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE OF 1227-1496 FROM N.A.
MEDLINE=85289337; PubMed=2411731;
Myers J.C., Loidl H.R., Seyer J.M., Dion A.S.;
"Complete primary structure of the human alpha 2 type V procollagen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  "Genetic distance of two fibrillar collagen loci, COL3A1 and COL5A2, located on the long arm of human chromosome 2."; Genomics 3:275-277(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE OF 1449-1496 FROM N.A.
MEDLINE-89138450; PubMed-3224983;
____sipourag_ P., Schwartz R.C., Liddell A.C., Salkeld C.S., Weil D.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    "Isolation of the alpha 3-chain of human type V collagen and characterization by partial sequencing.";
Biol. Chem. Hoppe-Seyler 373:69-75(1992).
                                                                                                                                                                                                                                                                                                                                                                               2178 RP--PSDLYLPPPDHGAPARGSP---HSEGGKRSPEPNKTSV 2214
                                                                                                                                                                                                                                                                                                                                                                                                                  1262 HPEWKSGDYWIDPNQGCTLDAMKVFCNMETGETCVYPNPATV 1303
                                                                                1093 SGEQGERGLKGHRGFTGLQGLPGPPGPSGDQGASGPAGPSGPR---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-JAPR-1988 (Rel. 07, Created)
01-JAN-1990 (Rel. 13, Last sequence update)
10-OCT-2003 (Rel. 42, Last annotation update)
Collagen alpha 2(V) chain precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1496 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COOH-terminal propeptide.";
J. Biol. Chem. 260:11216-11222(1985)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TISSUE=Placenta;
MEDLINE=92239022; PubMed=1571108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE OF 208-227.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CA25 HUMAN
P05997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ramirez F.;
                                            1955
                                                                                                                                                                                                                                                                                                                                       1209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Mann K.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PACE OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO OCCOOCO O
                                        8
                                                                                셤
                                                                                                                                                                g
                                                                                                                          à
                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                  ద
                                                                                                                                                                                                                                                                                                                                                                                 ò
```

<i>ት</i> ብ	1315	EGRVCRAISSASIEGLMGRAIPPERHSPHHLKEQHHIRGSITQGIPRSYVBAQEDYLRRE 1374
ò		REGTP-PPPPPSRDLTEAYKTQALGPLKLKPAHEGLVATVKEAGR 1
쉽	7	
ò	1424	LRHTPELPLAPRPLKEGSITQGTPLKYDTGASTTGSKKHDVRSL
Q	721	- - - - - - - - - - - - - -
ò	1481	GRIFPPVHPLDVMADARALERACYEESLKSRPGTASSSGGSIARGAPVIVPELGKPRQSP 1540
а	750	GDTGPPGLQGMPGERGIAGTPGPKGDRGGIGEKGAEGTA 788
ò	1541	LTYEDHGAPFAGHLPRGSPVTWREPTPRLQEGSLSSSKASQDRKLTSTPREIAKSPHSTV 1600
đ	789	GNDGAGGLPGPLGPPGPAGLLGEKGEPGPRGLVGPPGS 826
ò	1601	PEHHPHPISPYEHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLDAAAAYYLPRHLAPNPT 1660
q	827	RGNPGSRGENGPTGAVGFAGPQGSDGQPG 855
ò	1661	YPHLYPPYLIRGYPDTAALENRQTIINDYITSQQMHHNTATAMAQRADMLRGLSPRESSL 1720
Д	856	VKGBPGEP
ò	1721	ALNYAAGPRGIIDLSQVPHLPVLVPPTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPG 1780
qa	873	GPQGLAG-SPGPHGPNGVPGLKGGRGTQGPPGATGFPGSAGRVGPP 917
ò	1781	GPTHLTKPTTTSSSERERDRDRERDRDREREKSILTSTTTVEHAPIWRPGTEQSSGSSGS 1840
qq	918	GPAGAPGPAGPLGEPGKEGPPGPRGD 943
ò	1841	TODALQOR
qq	944	PGSHGRVGVRGPAGPPGGPGDKGDPGEDGQPGPDGPPGPAGTTGQRGI 991
ò	1887	ITAVEPSKPTVLRSTSTSSPVRPAATFPPATHCPLGGTLDGVYP 1930
q	992	-B
ò	1931	TLMEPULLPKEAPRVARPERPRADTGHAFLAKPPARSGLEPASSPSKGSEPR 1982
qa	1042	EP
ò	1983	PLVP-PVSGHATIPRTPAKNLAPHHASPDPPAPASASDPH 2022
QQ	1082	PGTPGPVGAPGDAGQRGDPGSRGPIGHLGRAGKRGLPGPQGPRGDKGDKGDRGDKGVKH 1141
ò	2023	REKTQSKPFSIQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDKGLPKHL 2074
qq	1142	RGFTGLÖGLFGPPGPNGEQGSAGIPG-PFGPRGPFGPV-GPSGKEGNPGPL 1190
δ	2075	EELDKSHLEGELRPKQPGPVKLGGEAAHLPHLRPLPE 2111
g	1191	GPLGPPGVRGSVGEAGPEGPPGEPGPPGPPGHLTAALGDIMGHYDESMPDPLPEFTE 1250
ð	2112	SQPSSSPLLQTAPGVKGHQRVVTLAQHISEVITQDYTRHHPQQLSAPLPAPLYSFPGASC 2171
ф	1251	DQAAPDDKNKTDPGVHATLKSLSSQIETMRSPDGSKKHPARTC 1293
δ.	2172	1666
qq	1294	DDLKLCHSAKQSGEYWIDPNQGSVEDAIK 1322
RESULT UN89 CA ID UN	95 NEEL 189 CAE	EL STANDARD; PRT; 6632 AA. Q17362;

```
PROSITE; PS50002; SH3; 1.
Muscle protein; Immunoglobulin domain; Repeat; SH3 domain;
                                                                     SH3.

DH.

H.

IG-LIKE C2-TYPE 1.

IG-LIKE C2-TYPE 2.

IG-LIKE C2-TYPE 3.

IG-LIKE C2-TYPE 4.

IG-LIKE C2-TYPE 4.

IG-LIKE C2-TYPE 4.

IG-LIKE C2-TYPE 5.

IG-LIKE C2-TYPE 6.

THR-RICH.

RCSD 3.

RCSD 4.

RCSD 4.

RCSD 6.

IG-LIKE C2-TYPE 10.

IG-LIKE C2-TYPE 10.

IG-LIKE C2-TYPE 11.

IG-LIKE C2-TYPE 11.

IG-LIKE C2-TYPE 11.

IG-LIKE C2-TYPE 12.

IG-LIKE C2-TYPE 13.

IG-LIKE C2-TYPE 13.

IG-LIKE C2-TYPE 14.

IG-LIKE C2-TYPE 19.

IG-LIKE C2-TYPE 19.

IG-LIKE C2-TYPE 19.

IG-LIKE C2-TYPE 20.

IG-LIKE C2-TYPE 20.

IG-LIKE C2-TYPE 21.

IG-LIKE C2-TYPE 21.

IG-LIKE C2-TYPE 21.

IG-LIKE C2-TYPE 21.

IG-LIKE C2-TYPE 21.

IG-LIKE C2-TYPE 21.

IG-LIKE C2-TYPE 31.

IG-LIKE C2-TYPE 31.

IG-LIKE C2-TYPE 31.

IG-LIKE C2-TYPE 31.

IG-LIKE C2-TYPE 31.

IG-LIKE C2-TYPE 31.

IG-LIKE C2-TYPE 31.

IG-LIKE C2-TYPE 31.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C2-TYPE 41.

IG-LIKE C3-TYPE 41.

IG-LIKE C3-TYPE 41.

IG-LIKE C3-TYPE 41.

IG-LIKE C3-TYPE 41.

IG-LIKE C3-TYPE 41.

IG-LIKE C3-TYPE 41.

IG-LIKE C3-TYPE 41.

IG-LIKE C3-TYPE 41.

IG-LIKE C3-TYPE 41.

IG-LIKE C3-TYPE 41.

IG-LIKE C3-TYPE 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL.
POTENTIAL.
POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL.
POTENTIAL.
POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL.
POTENTIAL.
POTENTIAL.
                                                                       3127

4930

633

633

633

633

11033

11223

1223

1223

1223

1223

1223

1223

1223

1223

1223

1233

1233

1233

1233

1233

1233

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

1333

13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          4485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              4678
4771
5057
5057
5057
5260
5366
5578
5578
6014
6014
60130
6239
6368
6368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              621
2975
3065
3759
3890
                                                                       4588
4681
4873
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5487
5595
5701
5815
5925
6038
6150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       6413
6507
568
2908
                                                  3D-structure.
DOMAIN 15
DOMAIN 14
DOMAIN 34
DOMAIN 54
DOMAIN 64
DOMAIN 74
DOWAIN 94
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DISULFID
DISULFID
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DISULFID
DISULFID
DISULFID
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DISULFID
                                                                                                                                                                                                                                                        DOMAIN
DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
               This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its war by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial or send an email to license@isb-sib.ch).
                                                                                                                                                                                 SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.

STRAIN-Bristol N2;

MEDLINE-961802778; PubMed=8603916;

Benian G.M. Tinley T.L., Tang X., Borodovsky M.;

"The Caenorhabditis elegans gene unc-89, required for muscle M-line assembly, encodes a giant modular protein composed of Ig and signal transduction domains.";

"I can seembly, encodes a giant modular protein composed of Ig and signal transduction admains.";

"I can laio. 132:835-848(1996).

SEQUENCE FROM N.A.

SEQUENCE FROM N.A.

STRAIN-Bristol N2;

Du Z., Le T.T., Wilson R.;

Submitted (WAY-1997) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Waterston R.;
Submitted (ApR-2002) to the EMBL/GenBank/DDBJ databases.
-!- FUNCTION: Structural component of the muscle M-line. Myofilament
lattice assembly begins with positional cues laid down in the
basement membrane and muscle cell membrane. UNC-89 responds to
these signals, localizes, and then participates in assembling an
                                                                                                               Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
Rhabditidae; Peloderinae; Caenorhabditis.
POPI_TaxID=6239;
10-OCT-2003 (Rel. 42, Created)
10-OCT-2003 (Rel. 42, Last sequence update)
10-OCT-2003 (Rel. 42, Last annotation update)
Muscle M-line assembly protein unc-89 (Uncoordinated protein 89)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IRR008957; FN III-like.
InterPro; IRR008957; FN III-like.
InterPro; IRR003961; FN III-like.
InterPro; IRR003108; Ig_c1;
InterPro; IRR001308; Ig_c2;
InterPro; IRR001849; PH.
InterPro; IRR001849; PH.
InterPro; IRR001849; PH.
InterPro; IRR001849; PH.
InterPro; IRR001849; PH.
InterPro; IRR001849; If.
InterPro; IRR001849; If.
Pfam; PR00041; Ig; 47.
Pfam; PR00041; Ig; 47.
Pfam; PR00189; H; 1.
Pfam; PR00189; H; 1.
Pfam; PR00189; IGC2; 23.
SWART; SW00408; IGC2; 23.
SWART; SW00326; SH3; 1.
PROSITE; PSS0010; DH 2; 1.
PROSITE; PSS0010; DH 2; 1.
PROSITE; PSS0003; PH_DOMAIN; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; U33058; AAB00542.1; -.
EMBL; AF003131; AAB54132.2; -.
PDB; 1FHO; 20-DEC-00.
WormPep; C09DI.1; CB30426.
                                                                                              Caenorhabditis elegans.
                                                                                                                                                                                                                                                                                                                                                                                                                                                 REVISIONS
```

636 AIARWGSKTVSQCKNF 1522SRT 686 APAAASEBAAFPPVVDD 1566 SPEKVEBKPASP 746 SDTESIPSPHTEAKDT 746 SDTESIPSPHTEAKDT 806 PTPPPAPSPSPEKSVVBEL 806 PTPPPAPSPSPEKSVVBEL 806 PTPPPAPSPSPEKSVVBEL 806 PTPPPAPSPSPEKSVVBEL 806 PTPPPAPSPSPEKSVVBEL 807 PTPPAPSPSPEKSVVBEL 807 PTPPAPSPSPEKSVVBEL 808 PTPPAPSPSPEKSVVBEL 809 PTPPAPSPSPEKSVVBEL 809 PTPPAPSPSPEKSVVBEL 809 PTPPAPSPSPEKSVVBEL 809 PTPPAPSPSPEKSVVBEL 809 PTPPAPSPSPEKSVVBEL 809 PTPPAPSPSPEKSVVBECKPTS 809 PTPPAPSPSPEKSVVBECKPTS 809 PTPPAPSPSPEKSVVBECKPTS 809 PTPPAPSPSPEKSVVBECKPTS		1846 -EKSPEKP 1099 PPTISNPPLISSAKHP	1270YBG-KKGHVLSY : : 2006 LHSPLYKWYQNGNLL 1324 SASIEGLMGRAIPPERH : : 2063 SAQLTVNPPST' 1384 PPPPPPSI 2102 PAPTVKWIINEKIVEESI 1432 ELRHTPELPLAPRPLI 2161 DIKVERNVKAPRESONI	
6 6 6 6 6 6 6 6 6	8 8 8 8 8	8 6 8 6 8	6 6 6 6 6 6 6	8 & 8 % 8 & 8
				
FT DISULPID 6946 5998 POTENTIAL. FT DISULPID 6036 6171 POTENTIAL. FT CONFLICT 2137 A -> P (IN REF. 1). FT CONFLICT 2245 2247 AKA -> P (IN REF. 1). FT CONFLICT 2284 2284 E -> P (IN REF. 1). FT CONFLICT 2284 E -> P (IN REF. 1). FT CONFLICT 2297 2297 M -> I (IN REF. 1). FT CONFLICT 3894 3888 DAGEY -> RRRH (IN REF. 1). FT CONFLICT 3929 3929 A -> P (IN REF. 1). FT CONFLICT 314 5134 A -> P (IN REF. 1). FT CONFLICT 5145 5145 T -> S (IN REF. 1). FT CONFLICT 5145 5145 T -> S (IN REF. 1). FT CONFLICT 5145 5145 T -> S (IN REF. 1). FT CONFLICT 5145 5145 T -> S (IN REF. 1). FT CONFLICT 5120 5120 L -> F (IN REF. 1). FT CONFLICT 5223 622 L -> F (IN REF. 1). FT CONFLICT 523 523 R -> N (IN REF. 1). FT CONFLICT 5245 C -> F (IN REF. 1). FT CONFLICT 5245 C -> F (IN REF. 1). FT CONFLICT 5245 C -> F (IN REF. 1). FT CONFLICT 5245 C -> F (IN REF. 1). FT CONFLICT 5245 C -> F (IN REF. 1). FT CONFLICT 5245 C -> F (IN REF. 1). FT CONFLICT 5245 C -> F (IN REF. 1). FT CONFLICT 5245 C -> F (IN REF. 1). FT CONFLICT 5245 C -> F (IN REF. 1). FT CONFLICT 5245 C -> F (IN REF. 1).	Duery Match 2.4%; Score 315; DB 1; Le Best Local Similarity 18.3%; Pred. No. 0.0068; Matches 318; Conservative 264; Mismatches 682; 22 HSLSYPVQIARTHTDVGLLEYQHHSRDYASHLSPG 896 HDGFYRISIPNTSMDKHDGEIVAKAQNEHGTAESRARLTVB 76 NERSQELHLRPB	956 TVKTGEFAVPETTVRCNPNPEVTWFINGHKMDQG 103IESKRPRLELLPDPLLRPSFLLAT 1016 LCRAENAVGRFETKARLVVLAPEKQKP 148 VSPPSPPHTDPELELVPPRLSKEELIQNMDRVDTVTWYLKGEELKQS-DRVDI 1068 EGEPKPTVTWYLKGEELKQS-DRVDI 195 KKQQQLEEBAARPPEPEKPVSPPIESKHRSLVQ 1119 TNSECSDETKAKITVOKKPPADFRPILP-	255 PLYNOPSDEIGANGINGARKKELLYFRENHARKOWKOKPCORY 30 1169 PTYEWSVIGRKVRDGDEGARVTRDESTUDGASILTIDTATYSSEUNHLTI 12 303 DQLMEALEKKVERIENNPRRAKESKVREYYEKOPPEIRKORELGERMOSRVGGRGS 35 1219 SVVAENTLGAEETGAQLTIEPKKESVVEKOPPEIRKORELGERMOSRVGGRGS 35 1219 SVVAENTLGAEETGAQLTIEPKKESVVEKOPPEIRKORELGERMOSRVGGRGS 35 360 GLSMSAARSHEVSEIIDGLSSEORN-GVESIAGQVKEASPEATT 12 360 GLSMSAARSHEVSEIIDGLSSEORN-EKOMRQLAVIPPMLYDADQORIKFIN 41 1274 TITMETSLTSTKTTTMSTTEVTSTVGGVTVETKESESESBATTVTGG	Db 1321 SGGVTEGSISVSKIEVVŠKTDSQTDVREGTPURARVSFABEELPKEVI 1367 Qy 471 YLTKKNENYKSLVRRSYRRGKSQQQQQQQQQQQQQQQQQQQQQQQPPRSSQEEKDEKEKEA- 529 Db 1368DSDRKKKKSPSPDKKEKSPEKTBERPASPTKKTGEEVKSPKEKSPASP 1415 Qy 530 EKEEEXPEVENDKEDLLKEKTDDTSGEDNDEKEAVASKGRKTANSQ 575
			•	

ò	636	AIARMVGSKTVSQCKNFYFNYKKRQNLDEILQQHKLKMEKERNARRKKKK 685
qq	1522	
ò	989	APAAASEEAAFPPVVEDEEMEASGVSGNEEEMVEEAEALHASGNEVPRGECSGPATVNNS 745
qq	1566	SPEKVEEKPASPKKEKSPEKPASPTKKSENEVK 1599
ò	746	PTPPR
б	1600	SPTKKEKSPEKSVVEELKSPKEKSPEKADDKPKSPTKKEKSPEKSATEDVKS 1651
ò	908	EEETA
g	1652	PIKKEKSPEKVEEKPISPIKKESSPIKKIDDEVKSPIKKEKSPQTVEEKPASPIK 1706
ò	864	SECTEEAEEGPAKGKDAEAAABAALKAEKKEGGSGRATTAKSSGAPQDSDSSA 919
qq	1707	KEKSPEKSVVEEVKSPKEKSPEKAEEKPKSPTKKEKSPEKSAAEEVKSPTKKEKSP 1762
ò	920	TCSADEVDEAEGGDKNRLLSPRPSLLTPTGDFRANASPQKPLDLKQLKQRAAAIPPIQVT 979
g G	1763	EKSAEEKPKSPTKKESSPVKMADD 1786
ò	086	KVHEPPREDAAPTK-PAPPAPPPQNLQPESDAPQQPGSSPRGKSRSPADPADKEAFAAE 1038
qq	1787	EVKSPTKKEKSPEKVEEKPASPTKKEKTPEKSAAEEL-KSPTKKEKSPSSPTKKTGDESK 1845
ò	1039	AOKLPGDPPCWTSGLPFPVPPREVIKASPHAPDPSAFSYAPPGHPLPLGLHDTARPVLPR 1098
qq	1846	
ò	1099	PPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEHAKAPVGPVTMGLPLPPMDP 1158
qq	1882	PPAPKLTRDLKLQTVNKTDLAHFEVVVEHATECKWFLDG 1920
ò	1159	KKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEASVLRGTALGSVPGGSITKGIPST 1215
qq	1921	KEITTAQGVTVSKDDQFEFR1955
ò	1216	RVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPSRLDRGREDSLPKGHVI 1269
qq	1956	- VASNAAGSVETKTELKVLETPKETKKPEFTDKLRDMEVTKGDTVQMDVIA 2005
ò	1270	YEG-KKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETAAPKRTYDMMEGRVGRAIS 1323
qq	2006	LHSPLYKWYQNGNLLEDGKNGVTIKNEENKSSLIIPNAQDSGKITVE-ASNEVGSSES 2062
ò	1324	HLKEQHHIRGS
g	2063	SAQLTVNPPSTTPIVVDGPKSVTIKETETAEFKATISGF 2101
ò	1384	PPPPPPSRDLTEAYKTQALGPLKLKPAHEGLVATVKEAGRSIHEIPRE 1431
đ	2102	PAPTVKWTINEKIVEESRTIT-TIKTEDVYTLKISNAKIEQTGTVKVTAQNSAGQDSKQA 2160
ò	1432	ELRHTPELPLAPRPLKEGSITQGTPLKYDTGASTTGSKKHDVRSLIGSPGRTFPPV 1487
g	2161	DIKVEPNVKAPKFKSQLTDKVADEGEPLRWNLELDGPSPGTEVSWLLNGQPLTKS 2215
ò	1488	HPLDVMADARALERACYEESLKSRPGTASSSGGSIARGAPVIVPELGKPR 1537
qq	2216	DTVQVVDHGDGTYHVTIAEAKPEMSGTLTAKAKNAAGECETSAKVTVNGGNKKPEFVQAP 2275
8.5	1538	QSPLTYEDHGAPFAGHLPRGSPVTWREPTPRLQEGSLSSSKASQDRKLTSTPREIAKSPH 1597
ò	- 6	
;	,	_

```
MUCIL HUMAN STANDARD; PRT; 1255 AA.
P15941; P13911; P15942; P17626; Q14128; Q14876; Q16437; Q16442;
Q16615; Q9BXA4; Q9BT57; Q9UEN6; Q9UQL1; Q9Y4J2;
Q1-JAN-1990 (Rel. 13, Created)
Q1-APR-1990 (Rel. 14, Lat sequence update)
Q1-APR-1990 (Rel. 42, Last annotation update)
Mucin 1 precursor (MUC-1) (Polymorphic epithelial mucin) (PEMT)
(Episialin) (Tumor-associated mucin) (Grcinoma-associated mucin)
(Tumor-associated epithelial membrane antigen) (EMA) (H22AG) (Peanut-areactive unitary mucin) (Broast carcinoma-associated antigen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A. (ISOFORM 5).

TISSUE=Breast carcinoma;

MEDLINE=9057413; PubMed=2351132;

Wreschner D.H., Hareuveni M., Tsarfaty I., Smorodinsky N., Horev J.,

Zaretsky J., Kotkes P., Weiss M., Lathe R., Dion A., Keydar I.;

"Huann epithelial tumor antigen cDNA sequences. Differential splicing
may generate multiple protein forms.";

Bur. J. Blochem. 189:463-473(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TISSUB-Breast carcinoma;
MEDLINE=90276414; PubMed=2112460;
Hareuveni M., Tsarfaty I., Zaretsky J., Kotkes P., Horev J.,
Zrihan S., Weiss M., Green S., Lathe R., Keydar I., Wreschner D.H.;
"A transcribed gene, containing a variable number of tandem repeats, codes for a human epithelial tumor antigen. cDNA cloning, expression of the transfected gene and over-expression in breast cancer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SECUENCE FROM N.A. (ISOFORM 1).

MEDLINE-91097524; PubMed=2268309;
Lancaster C.A., Peat N. Dubig T., Wilson D.,

Taylor-Papadimitriou J., Gendler S.J.;

"Structure and expression of the human polymorphic epithelial mucin Biochem. Biophys. Res. Commun. 173:1019-1029(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Lan M.S., Batra S.K., Qi W.-N., Metzgar R.S., Hollingsworth M.A., "Cloning and sequencing of a human pancreatic tumor mucin cDNA.", J. Biol. Chem. 265:15294-15299(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ligtenberg M.J.L., Vos H.L., Gennissen A.M.C., Hilkens J.;
"Episialin, a carcinoma-associated mucin, is generated by a
polymorphic gene encoding splice variants with alternative amino
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gendler S.J., Lancaster C.A., Taylor-Papadimitriou J., Duhig T., Peat N., Burchell J., Pemberton L., Lalani B.-N., Wilson D.; "Molecular cloning and expression of human tumor-associated J. Phymorphic epithelial mucin.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Biol. Chem. 265:15286-15293(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    termini.";
J. Biol. Chem. 265:5573-5578(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TISSUE=Pancreas;
MEDLINE=90368716; PubMed=2394722;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE=90202794; PubMed=2318825;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Eur. J. Biochem. 189:475-486(1990)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TISSUE=Breast carcinoma;
MEDLINE=90368715; PubMed=1697589;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A. (ISOFORM 1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A. (ISOFORM 1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A. (ISOFORM 1)
                                                                                                                                                                                                                                                                                                                                                            (CD227 antigen)
                                                                                                                                                                                                                                                                                                                                                                                                                                 Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      tissue.
                                                                            A PRESENTATION OF STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STANKING STA
```

MEDLINE=91033045; PubMed=1688329;

```
"Comparison of MUC-1 mucin expression in epithelial and non-epithelial cancer cell lines and demonstration of a new short variant form
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Zhang L.X., Li C.H., Sun L.Y., Yue W.; "Cloning of a new potential secreted short variant form of MUC1 mucin in epithelial cancer cell line."; Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
                                                                                                            SEQUENCE FROM N.A. (ISOFORM 7).

MEDLINE=95010060; PubMed=7925397;
Zrihan-Licht S., Vos H.L., Baruch A., Elroy-Stein O., Sagiv D.,
Zrihan-Licht and molecular cloning of a novel WUC1 protein,
"Characterization and molecular cloning of a novel WUC1 protein,
devoid of tandem repeats, expressed in human breast cancer tissue.";
Eur. J. Biochem. 224:787-795 (1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                "A highly immunogenic region of a human polymorphic epithelial mucin expressed by carcinomas is made up of tandem repeats."; J. Biol. Chem. 263:12820-12823(1988).
Tsarfaty I., Hareuveni M., Horev J., Zaretsky J., Weiss M., Jeltsch J.M., Garnier J.M., Lathe R., Keydar I., Wreschner D.H., "Isolation and characterization of an expressed hypervariable gene coding for a breast-cancer-associated antigen.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TISSUE-Lung;
MEDLINE-96181716; PubMed=8604237;
Yu C.J., Yang P.C., Shew J.Y., Hong T.M., Yang S.C., Lee Y.C., Tee L.N., Luh K.T., Wu C.W.;
Tee L.N., Luh K.T., Wu C.W.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PARTIAL SEQUENCE FROM N.A.
MEDLINE-88330762; PubMed-3417635;
Gendler S.J., Taylor-Papadimitriou J., Duhig T., Rothbard J.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE OF 1-160 FROM N.A. (ISOFORM 2).
MEDLINE=90088473; Pubmed=2597151;
Abe M., Siddiqui J., Kufe D.;
"Sequence analysis of the 5' region of the human DF3 breast carcinoma-associated antigen gene.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Weiss M., Baruch A., Keydar I., Wreschner D.H.;
"Preoperative diagnosis of thyroid papillary carcinoma by re
transcriptase polymerase chain reaction of the MUC1 gene.";
Int. J. Cancer 66:55-59(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Coombes R.C.;
                                                                                                                                                                                                                                                                               MEDLINE=97355747; PubMed=9212228;
Oosterkamp H.M., Scheiner L., Stefanova M.C., Lloyd K.O.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Zhang L.X., Li C.H.;
"Molecular cloning of an isoform of MUC1, MUC1/Y.";
Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Buluwela L., Liu Q., Lugmani Y.A., Gomm J.J., Coombes R.(
Submitted (OCT-1992) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Biochem. Biophys. Res. Commun. 165:644-649(1989)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE OF 1-46 FROM N.A. (ISOFORMS 3 AND 4). TISSUE=Breast carcinoma;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYDRATE-LINKAGE SITES WITHIN THE REPEAT.
                                                                                                                                                                                                                                                               SEQUENCE FROM N.A. (ISOFORMS 6; 7 AND 8).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE OF 1-109 FROM N.A. (ISOFORM 2)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE=96183746; PubMed=8608966;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A. (ISOFORM 9).
TISSUE=Epithelial cancer;
                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A. (ISOFORM 7).
                                                                                                                                                                                                                                                                                                                                                                                                 J. Cancer 72:87-94(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE OF 1-89 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Oncology 53:118-126(1996)
                                                                          Gene 93:313-318(1990)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TISSUE=Thyroid;
                                                                                                                                                                                                                                                                                                                     Finstad C.L.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Burchell J.;
```

Mon Apr

1812 KSILTSTTTVEHAPIWRPGTEQSSGSSGGGGGGSS--SRPASHSHAHQHSPISPRTQD 1869 1163 PFSGVKQEQLSPRGQAGPPESLGVPTAQEASVLRGTALGSVPGGSITKGIPSTR-VPSDS 1221 1222 AITYRGSITHGTPADVLYKGTITRIIGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYE 1281 1282 GCMSVTQCSKEDGRSSSGPPHETAAPKRTYDMMEGRVGRAISSASIEGLMGRAIPPERHS 1341 1342 PHHLKEQHHIRGSITQGIPRSYVEAQEDYLRREAKLLKREGTPPPP----PPSRDLTEAY 1397 1398 KTQALGPLKLKPAHEGLVATVKEAGRSIHEIPREELRHTPELPLAPRPLKEGSITQGTPL 1457 1458 KYDTGASTTGSKKHDVRSLIG---SPGRTFPPVHPLDVMADARALERACYEESLKSRPGT 1514 1515 ASSSGGSIARGAPVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSPVTMREPTPRLQEGSL 1574 1695 MHHNTATAMAQRADMLRGLSPRESSLALNYAAGPRGIIDLSQVPHLPVLVPPTPGTPATA 1754 1930 PTLMEPVLLPKEAPRVARPERPRADTGHAFLAKPPAR----SGLEPASSPSKGSEPRPLV 1985 1755 MDRLAYLP-TAPQPFSSRHSSSPLSPGGPTH--LTKPTTTSSSERERDRDRERDRDRERE 1811 610 478 648 925 EVDEAEGGDKANRLLSPRPSLLTPTGDPRANASPQKPLDLKQLKQRAAAIPPIQ-VTKVHE PGS------TAPPAHGVTSAPDT-RPAPGSTAPPAHGV-----TSAPDTRPAPGST 649 -GSTAPPAHGVTSAP-DTRPAPGSTAPPAHGVTSAPDTRPAPG-----STAPPAHGV 365 ECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSSGAPQDSDSSATCSAD ----HGVTSAPDTRPAPGSTAPPAHGVTSAPD -----TRPAPGSTAPPAHGVTSAPDTRPAP------GSTAPPAHGVT----984 PPREDAAPTKPAP-PAPPPPQNLQPESDAPQQPGSSPRGKSRSPAPPADKEAFAAEAQKL ----SAPDTRPAPGSTAPPAHGVTSAPDTRPAPGST-----APPAHGVTSAPDTRPA PGDPPCWTSGLPFPVPPREVIKASPHAPDPSAFSYAPPGHPLPLGLHDTARPVLPRPPTI ------PDTRPA 1575 SSSKASQDRKLTSTPREIAKSPHSTVPEHHPHPISPYEHLLRGVSGVDLYRSHIPLAFDP 1870 ALQQRPSVLHNTGMKGIITAVEPSKPTVLRSTSTSSPVRPAATFPPATHCPLGGTLDGVY SNPPPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEHAKAPVGPVTMGLPLPMDPKKLA P-----APPAHGVTSAPDTRPAPGST--APPAHGVTSA-PDTRPAPGST A----STAPPAHGVTSAPDTR----PAPG-----STAPPAHGV----------TRPAPGSTAPPAHGVTSAP -----GSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTR--------PAPGS 611 TAPPAHGVT-SAPDTRPAPGST-----APPA-HGVTSAPDT--RPAP----TSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAALENRQTIINDYITSQQ ----GSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTA----PPAHGVTS APDTRPAPGSTAPPAHGVTSAPDTRPA------PGSTAPPA-----HGV-775 AHGVTSAPDTRPAP----GSTAP----PAHGVTSAPDTRPAP------------APDTRPAP ---DTRPAPGSTAPPAHGVTSAPDT-----------GSTAP--PAHGVTSAPD-DTRPAPGSTAPPAH-GVTS-----------TSAPDTRPA---PGSTAPP-GVTSAPDTRPAPGSTAPPA 1043 1635 280 328 1103 408 452 479 508 543 569 743 g a 셤 g 셤 셤 ò ઠે 셤 g 원 셤 δ 요 셤 셤 ò g à 원 8 셤 ò ò ò 8 g à 원 ઠે ò 8 ò ò ઠે 8

```
----PPV-----SGHATIARTPAKNLAPHHASPDPPAPASASDPHREKTQSKPFSIQE 2035
                                                                                                                        2036 LELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDKGLPKHLEELDKSHLEGELRPKQPGPVK 2095
                                                                                                                                                                                                                                        1054
995
                                                                                    949 GSTAPPVHNVTSASGSAS---GSASTLVHNGTSARATTTPAS------KSTPFSIP-
                                                                                                                                                                                           2096 LGGEAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQRVVTLAQHISEVITQDYTRHHPQQL
                                                                                                                                                                                                                       Σ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE OF 1254-1712 FROM N.A.
MEDLINE=87219156; PubMed=3582677;
HOSTIKAS S.L., Kurkinen M., Tryggvason K.;
"Nucleotide sequence coding for the human type IV collagen alpha 2 chain cDNA reveals extensive homology with the NC-1 domain of alpha 1 (IV) but not with the collagenous domain or 3'-untranslated region.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Brazel D., Pollner R., Oberbaeumer I., Kuehn K., "Human basement membrane collagen (type IV). The amino acid sequence of the alpha 2(IV) chain and its comparison with the alpha 1(IV) chain reveals deletions in the alpha 1(IV) chain ";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE OF 1486-1712 FROM N.A.
MEDLINE=87250571; PubMed=2439508;
Myers J.C., Howard P.S., Jelen A.M., Dion A.S., Macarak E.J.;
Myers J.C., Howard P.S., Jelen A.M., Dion A.S., Macarak E.J.;
Eupplication of type IV collagen COOH-terminal repeats and species-
specific expression of alpha 1(IV) and alpha 2(IV) collagen genes.";
J. Biol. Chem. 262:9231-9238(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2 chain of human type chain.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE OF 1451-1485 FROM N.A.
MEDLINE=87092438; PubMed=3025878;
Griffin C.A., Emanuel B.S., Hansen J.R., Cavenee W.K., Myers J.C.,
"Human collagen genes encoding basement membrane alpha 1 (IV) and
alpha 2 (IV) Chains map to the distal long arm of chromosome 13.",
Proc. Natl. Acad. Sci. U.S.A. 84:512-516(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                         ----SHH--SDTPTTLASHSTKTDASSTHHSSVP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                   , Last sequence update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hostikka S.L., Tryggvason K.,
"The complete primary structure of the alpha
collagen and comparison with the alpha 1(IV)
J. Biol. Chem. 263:19488-19493(1988).
                                                                                                                                                                                                                                                                  SAPLPAPLYSFPGASCPVLDLRRPPSDLYL 2185
                                                                                                                                                                                                                                                                                           1055 NSSLEDPSTDY-----YQELQRDISEMFL 1078
                                                                                                                                                                                                                                                                                                                                                                                                                            01.AUG-1988 (Rel. 08, Created)
01-F8B-1991 (Rel. 17, Last sequence u
16-0CT-2001 (Rel. 40, Last annotation
Collagen alpha 2(IV) chain precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
MEDLINE-89066769; Pubmed=3198637;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE OF 1-1042 FROM N.A.
TISSUE-Placenta;
MEDLINE-88151998; PubMed=3345760;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Biochem. 172:35-42(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FEBS Lett. 216:281-286(1987)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE OF 1-33 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                            HUMAN
                  899
                                                    1986
                                                                                                                                                                                                                                 1024
                                                                                                                                                                                                                                                                  2156
                                                                                                                                                             966
                                                                                                                                                                                                                                                                                                                                                                         CA24_HUMAN
                                                                                                                                                                                                                                                                                                                                                          RESULT
                                                  à
                                                                                    유
                                                                                                                        à
                                                                                                                                                      요
                                                                                                                                                                                             ò
                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                8
```

```
nidogen.

-!- SUBNIT: There are six type IV collagen isoforms, alpha 1(IV)-
alpha 6(IV), each of which can form a triple helix structure
alpha 6(IV), each of which can form a triple helix structure
with 2 other chains to generate type IV collagen network.
C -!- DOWAIN: Alpha chains of type IV collagen have a noncollagenous
domain (NCI) at their C-terminus, frequent interruptions of the
G-X-Y repeats in the long central triple-helical domain (which may
criple-helical TS domain.
C triple-helical TS domain.
C triple-helical TS domain.
C -!- PTM: Prolines at the third position of the tripeptide repeating
unit (G-X-Y) are hydroxylated in some or all of the chains.
C -!- PTM: Type IV collagens contain numerous cysteine residues which
are involved in inter- and intramolecular disulfide bonding. 12 of
these, located in the NCI domain, are conserved in all known type
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                       Soininen R., Huotari M., Hostikka S.L., Prockop D.J., Tryggvason K.; "The structural genes for alpha 1 and alpha 2 chains of human type IV collagen are divergently encoded on opposite DNA strands and have an
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TISSUE-Placenta,
MEDLINE-89005112; PubMed=2844531;
Siebold B., Deutzmann R., Kuehn K.;
Siebold B., Deutzmann R., Kuehn K.;
Siebold B., Deutzmann R., Kuehn K.;
"The arrangement of intra- and intermolecular disulfide bonds in the carboxyterminal, non-collagenous aggregation and cross-linking domain of basement emembrane type IV collagen.";
Eur. J. Biochem. 176:617-624(1988).
-!- FUNCTION: Type IV collagen is the major structural component of glomerular basement membranes (GBM), forming a 'chicken-wire' meshwork together with laminins, proteoglycans and entactin/
                                                                                                                                                                                                                                                                                                                                                                                                                                                              οĘ
                                                                                                                                                                                                                                                                                                                                                                              MEDLINE=93305049; PubMed=8317999;
Fischer G., Schmidt C., Opitz J., Cully Z., Kuehn K., Poeschl E.;
"Identification of a novel sequence element in the common promoter
region of human collagen type IV genes, involved in the regulation
divergent transcription.";
                                                                                                                                                    SEQUENCE OF 1-33 FROM N.A.

MEDLINE=89030632; PubMed=2846280;
Poeschl E., Pollner R., Kuchn K.;

"The genes for the alpha 1(IV) and alpha 2(IV) chains of human basement membrane collagen type IV are arranged head-to-head and separated by a bidirectional promoter of unique structure.";

EMBO J. 7:2687-2695(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE OF 1480-1535; 1545-1614; 1617-1701 AND 1705-1712
                                                                                           overlapping promoter region.";
J. Biol. Chem. 263:17217-17220(1988)
       MEDLINE=89034231; PubMed=3182844;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Biochem. J. 292:687-695(1993).
                                                                                                                                                                                                                                                                                                                                      SEQUENCE OF 1-33 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                              TISSUE=Skin;
```

NAS.

MIM; 120090; -. GO: Coollagen type IV; TAS. GO: 0005587; C:collagen type IV; TAS. GO: 0005501; F:extracellular matrix structural constituent; GO: GO: 00050198; P:extracellular matrix organization and bioge. InterPro; IPR008161; Clg helix. InterPro; IPR008160; Collagen.

EMBL; X05562; CAA29076.1; -EMBL; X05610; CAA29098.1; -. EMBL; J02760; AAA58422.1; -. EMBL; M36963; AAA53099.1; -.

EMBL; X12784; CAA31275.1; -. EMBL; J04217; AAA53097.1; -. PIR; A32024; CGHU2B.

EMBL; EMBL; EMBL;

HGNC:2203; COL4A2.

Genew;

```
G-RAIPPERHSPHHLKEQHHIRGSITQGIPRSYVEAQEDYLRREAKLLKREGTPPP---- 1386
                                                                                                                                                                                                                                                                                                                                                                                                                                                           HPLDVMADARALERACYEESLKSRPGTASSSGGSIARGAPVIVPELGKPRQSPLTYEDHG 1547
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ESSLALNYAAGPRGIIDLSQVPHLPVLVPPTPGTPATAMDR-LAYLPTAPQPFSSRHSSS 1775
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AGLYGEIGATGDFG--DIGDTINL----PGRPGLKGERGTTGIPGLKGFFGEKGTEG 1128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GTEOSSGSSGSSG-----GGGGSSSRPASHSHAHQHSPISPRIQDALQQRPSVLHNTGMK 1884
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GLTGPPGSQGELGRIGLPGGKGDDGWPGA---------PGLPGFPGLR 1193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GII-----TAVEPSKPTVLRSTSTSSPVRPAATFPPATHCPLGGTLD-GVYPTLMEPV 1936
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LLPKEAPRVARPERPRADTGHAFLAKPPARSGLEPASSPSKGSEPRPLVPPVSGHATIAR 1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -----FPGITPPSNISG 1278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GDRGPKG-PKGDPGFPGAPGTVGAPGIAGIPQKIA---IQPGT-VGPQGRRGPPGA---- 1415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GLPGFPGVAGPPGITGFPG---------FIGSRGD--KGAPGR 1077
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ---Irgvravgapparkagrafp 1154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GİRGIHGLPGİKGFPGSP-----GSDIHGDPGFPG----PPGERGDPGEANTLPGPV 1241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1997 TPAKNLAPH----HASPDPPAPPASASDPHREKTQSKPFSIQELELRSLGYHGSSYSPEG 2052
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----GAPGTPGTKGW 1327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AGDSGPQGRPGVF---GLP-----GEKGPR------GEQGFMGNTGPTGAV 1364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ----PESQPSSSPLLQTAPGVKGHQRVVTLAQHISEVITQDYTRHHPQQLSAPLPAPLYS 2165
                                                                                                                                                                                                              ----GL------PGPDGSPGPIGLPGPDGPPGERGLPGEVLGAQPGPRG----DAGVPGQP 793
                                                                                                                                                                                                                                                                                                                 ----PPPSRDLTEAYKTQALGPLKLK--PAHEGLVATVKEAGRSIHBIPREELRHTPELP
                                                                                                                                                                                                                                                                                                                                                                                        1441 LAPRPL-KEGSITQCTPLKYDTG-ASTTGSKKH-----DVRSLIGSPG----RTFPPV
   608 IPGTKGTPGEMGPPGLGLPGLKGQRGFPGDAGLPGPPGFLGPPGPAGTPGQIDCDTDVKR
                                                                                                       DSAITYRGSITHGTPADVLYKGTITRIIGEDSPSRLDRGRED-SLPKGHVIYEGKKGHVL
                                                                                                                                                                           SYEGGMSVTQCSKEDGR-----SSSGPPHETAAPKRTYDMMEGRVGRAISSASIEGLM
                                                                                                                                                                                                                                                                                                                                                     -----PGD--RGDPGDT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1548 APFAGHLPRGSPVTMREPTPRL------OEGSLSSSKASQDRKLTSTPREIAKSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HSTVPEHHPHPISPYEHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLDAAAAYYLPRHLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1657 PNPTYPHLYPPYLIRGYPDTAALENRQTIINDYITSQQMHHNTATAMAQRADMLRGLSPR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1776 PLS-PGGPTHLTKPTTTSSSERERDRDRERDREREKSILTSTTTVEHAPIWR----P
                                    -FSGVKQEQLSPRGQAGP---PESLGVPTAQEASVLRGTALGSVPGGSITKGIPSTRVPS
                                                                    668 AVGGDRÓEAIOPGCIAGPKGLPGLPGPFGPTGAKGLRG----IPGFAGADGGPGPR---
                                                                                                                                         ------GLPGDA-----GRPGPPGFIGPRGSKGAV-
                                                                                                                                                                                                                                                                                  GLKGLPGDRGPP-----GFRGS--QGMPGM------PGLKGQPGLPGPSGQP
                                                                                                                                                                                                                                                                                                                                                                                                                      GAPGPVGMKG-----LSGDRGDAGFTGEQGHPGSPGFKGIDGMPGTPGLKGDRGSPGM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VEPVSPVSSPSLTHDKGLPKHLEELDKSHLEGELRPKQPGPVKLGGEAAHLPHLRPL---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 --GIPG-----LSGIPGLPGRPGHIKGVKGDIGVP-GIP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -GFKG--SRGDPGPPGPPVILPGMKDIKGEKGDEGPMGLKGYLGAKGIQGMP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DGFQGM------PGLKGRPGFPGSKGEAGFFGIPGLKGLAGEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GVPGQ-----KGDQGAPGERGPPGSPGLQG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1279 APGDKGAPGIFGLKGYRGPPGPPGSAALPGSKGDTGNP
                                                                                                                                                                                                                                                                                                                                                  GLYGPPGLHGFPGAPGQE--GPLGLPGIPGREGI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1129 DIGFPG------
                                                                                                                                                                                                                                                1332
                                                                                                                                                                                                                                                                                                                                                                                                                                                           1488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1830
                                                                                                                                                                                                                                                                                                                   1387
                                                                                                                                                                                                                                                                                                                                                                                                                        875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1717
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1885
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1937
                                  Š
                                                                  원
                                                                                              8
                                                                                                                                셤
                                                                                                                                                                          ò
                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                           유
                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                   음
                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            용
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   용
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    음/ 등
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          용
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ਨੇ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -AIPPIOVTKVHEPPREDAAPTKPAP----PAPPPP-----QNLQPESDAPQQPGS-SP 1019
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PKEEKEEETA; -----AAPPVEEGEEQK-------PPAAEELAVDTGKA 858
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    415 PGİPAL----YGGPPGPDGKRGPPGPPGPPGPPGPPGPLFGLKGAKGRAGFPGLPGSPGA 470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -PPATLGADGP-----PPGPPTPPRRTSRAPIEPTPASEATGAPTPPPAPPSPSAPPPVV 822
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ------PGSEGFTGPPGPQGPKGQKGEPYAL 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           235 FYGVKGEKGDVGQPGPNGIPSDTLHPIIAPTGVTFHPDQYKGEKGSEGEPGIRGISLKGE 294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PSLAKGARGDPGFPGAQGEPGSQGEPGDPGLPGPPGLSIGDGDQRRGLPGEMGPKGFIGD 414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RG-----PKGWKGDAGECRCTEGD------EAIKGLPGLPGLPGFAGIN 508
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ----AKGKDAEAAEATAEG-ALKAE 894
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             KK----EGGSGRATTAKSSGAPQDSDSSATCSADEVD----EAEGGDKN----RLLSPR 941
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EGIMGFPGLRGYPGLSGEKGSPGQKGSRGLDGYQGPDGPRGPKGEAGDPGPPGLPAYSPH 354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GVSGNEEEMVEEARALHASGNEVPRGECSGPATVNNSSDTESIPSPHTEAAKDTGQNGPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RGKSRSPAPPADKEAFAAEAQKLPGDPPCWTSGLPFPVPPREVIKASPHAPDPSAFS---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GEPGRKGDKGDPGQHGLPGPPGLKGVPGNIGAPGPKGAKGDSRTITTKGERGQPGVPGVP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PSILT------PTGDPRANASPQKP-----LDLKQLKQRAA------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GMKGDDGSPGRDGLDGFPGLPGPPGDGIKGPP---GDPGYPG-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ------YAPPGH-----PLPLGLHDTAR------
                                                                                                       matrix, Connective tissue, Repeat, Hydroxylation,
Basement membrane, Collagen, Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DB 1; Length 1712;
                                                                                                                                                                                                                                                                                                                                N-LINKED (GLCNAC. ..).
R -> P (IN REF. 2).
A -> G (IN REF. 2).
A -> I (IN REF. 5).
G -> H (IN REF. 5).
G -> H (IN REF. 9).
W, 2582A17847890037 CRC64;
                                                                                                                                                                    COLLAGEN ALPHA 2 (IV) CHAIN.
TRIPLE-HELICAL REGION.
NONHELICAL REGION (NCI).
OR 1590 (BY SIMILARITY).
OR 1593 (BY SIMILARITY).
OR 1705 (BY SIMILARITY).
OR 1706 (BY SIMILARITY).
OR 1708 (BY SIMILARITY).
OR 1708 (BY SIMILARITY).
OR 1708 (BY SIMILARITY).
A 1708 (GLCNAC. . .).
N-LINKED (GLCNAC. . .).
A -> G (IN REF. 2).
A -> G (IN REF. 2).
A -> G (IN REF. 9).
H -> G (IN REF. 9).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches 407; Conservative 146; Mismatches 704; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ---LPMDPKKLAP---
                                                                                                                                                       AMINO-TERMINAL PROPEPTIDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2.4%; Score 312.5; DB
9.8%; Pred. No. 0.0023
InterPro; IPR001442; Procollagn4_C.
                                             Probom; PD000007; Clg helix; 7.
Probom; PD003923; ProcollagnC4; 1.
SWART; SM00111; C4; 2.
Extracellular matrix; Connective t.
Glycoprotein; Basement membrane; C3
SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GRPGYDGCNGTQGDSGPQGP-
                                                                                                                                                                                                                                                                                                                                                                                                 1663 1663
1701 1701
1712 AA; 167535 M
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    VPYSEHAKAPVGPVTMGLP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           19.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EEPVKSECTEEAEEGP-
                                                                                                                                                                                                                            1593
1590
1555
1708
                                                                                                                                                                           1712
1484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity
                                                                                                                                                                                                                                                               549
                                                                                                                                                                                                                                            DISULFID
DISULFID
DISULFID
DISULFID
DISULFID
CARBOHYD
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1020
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              607
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        859
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             895
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                942
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1077
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             509
                                                                                                                                                                                                                            DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local
                                                                                                                                                                                        DOMAIN
                                                                                                                                                       PROPEP
 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8 6 8 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ર્જ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ठ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ર્જ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                용
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ઠે
```

•

```
(See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  139 RSLIGKLEPVSPPSPPHIDPELELVPPRLSKEELIQNMDRVDREITMVEQQISKLKKKQQ 198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              256
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | :: | : | : | 229 TVRSPSKLKNSSLDAPNATSPDLVVESPCPPCTSYPVHVASTQKSEQVAAQCYSHLY--- 285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SVVSQSVEQWCSLLL------RDQKPKKQGKYICEYCNRACAKPSVLLKHIRSHT--- 430
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     639 RMVGSKTVSQCKNFYFNYKKRQNLDEILQQHK-----LKMEKERNARRKKKKAPAAASE 692
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       199 QLEBERAKPPEPEKPVSP--PPIESKHRSLVQIIYDENRKKAEAAHRILEGLGPQVELPL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          --SASESHKKHNCVPAKOGRO-----PTKONGETPGMTAESSESGDLVS---PKKT---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               422 -VYKDRQVMNMWSEQEKETFREKFMQHP----KNFGL-IASFLERKTV-AECVLYYYLTK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            54 VAENHLKKI PK----SPLRNPLQTKHK-----QNTEEP-----PFSVLP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         257 YNQPSDTRQYHENIKINQAMRKKLILYFKRRNHARKQWKQKFCQRYDQLMEALEKKVERI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            317 ENNPRRRAKESKVREYYEKOFPEIRKORELQERMOSRVGORGSGLSMSAARSEHEVSEII
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          135 -SSPHQR---SELRRW------RSEGSDPTRLSGLDGQRDSSSSSKART-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              377 DGLSEQENLEKOMRQLAVIPP-----MLYDADQQRIKFINMNG----LMADPMK---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                175 ----- DNSECSSPCCSTTPPSYTSTAFDVLLKAMEPELSTLSQKGSSCAIKTEKLRPNK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -----SSQDHLVPKLSQQNQQLPGHLGFTGSLTNLHTLESTKLEPI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AITPOOSAELASMELNESSRWTEEEMETAKKG-------LLEHGRNWSAIA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   475 KNENYKSLVRRSYRRRGKSQQQQQQQQQQQQQQQQPMPRSSQEEKDEKEKEKEAEKEEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    535 KPEVENDKEDLLKEKTDDTSGEDNDEKEAVASKGRKTANSQGRRKGRITRSMANEANSEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           327 YNTAVTSTVGLTSPSTRTQVTPPHQQMDSVSPLSVSPASSTQSPPGPI-YSSAHVA----
                                                                                                                                                                                                                                                                          PROSITE; PS00028, ZINC FINGER C2H2 1; 4.
PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.
Transcription regulation; Zinc-finger; Metal-binding; DNA-binding;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2.4%; Score 311; DB 1; Length 2688; larity 18.2%; Pred. No. 0.0039; Conservative 376; Mismatches 1073; Indels 98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POLY-ASP.
W; 5EAD46C3A7008BE6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                C2HC-TYPE (POTENTIAL)
POLY-SER.
C2H2-TYPE.
C2H2-TYPE.
equires a license agreement (email to license@isb-sib.ch)
                                                                                                                                                                                                                                                                                                                                                            C2H2-TYPE.
C2H2-TYPE.
POLY-SER.
POLY-PRO.
                                              EMBL; L36829; AAA98810.1; -...
EMBL; L36825; AAA98810.1; JOINED.
EMBL; L36826; AAA98810.1; JOINED.
EMBL; L36827; AAA98810.1; JOINED.
EMBL; X68946; CAA48762.1; -...
PIR; L49477; 149477.
                                                                                                                                                                                                                                                                                                                                                                                                                                      76
1486
4 2096
2126
2182
; 288341 Mr
                                                                                                                                                                                                     MGD; MGI:96100; Hivepl.
InterPro; IPR007087; Znf C
Pfam; PF00096; zf-C2H2; 5.
                                                                                                                                                                                                                                                           C2H2;
                                                                                                                                                                                                                                                                                                                            Nuclear protein; Repeat
DOMAIN 165 170
                                                                                                                                                                                                                                                           SMART; SM00355; ZnF
                                                                                                                                                                        HSSP; P15822; 3ZNF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Similarity
                                                                                                                                                                                       TRANSFAC; T00007;
                                                                                                                                                                                                                                                                                                                                                                                                                                                 1482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Simi
Matches 540;
                   or send an
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                            ZN_FING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    286
                                                                                                                                                                                                                                                                                                                                                                                                                                                               ZN FING
                                                                                                                                                                                                                                                                                                                                                                                                                                ZN FING
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                     - PGEIGP----OGPPGE----PGFRGAPGKAGPQGRGG------VSAVPGFRGDEG-- 1456
                                                                                                                                                                                     1536
                                                                                                                                                                                                              AIIRKALM------GKYDQ--WEESP-----PLSANAFNPLNASASLPAAM 2381
     2166 PPGASCPVLDLRRPPSDLYLPPPDHGAPARGSPHSEGGKRSPEPNKTSVLGG--GEDGIE 2223
                                                                                                                                                                                                                                                                                    -----HTLTSPGGGKAKVSGRPSSRKAKSPAPG 2420
                                                                                                                                                                                                                                                                                                                          -----SPGSC 1646
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bicinformatics and the EMBL outstation-the European Bicinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial
                                                                                                                                          2284 KSKKQEINKKLNTHNRNEPEYNISQPGTEIFNMPAITGTGLMTYRSQAVQEHASTNMGLE
                                                                                                                                                                          1483 -GRSVSİGYLLVKHSQTDQE----PMCPVGMNKLWSGYSLLYFBGQEKAHNQDLGLAGS
                                                                                                                                                                                                                                       2224 PVSPPEGMTEPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQPPAFFSKLTESNSAMV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TISSUE=Lens epithelium;
MEDLINE=90281161; PubMed=1694016;
Nakamura T., Donovan D.M., Hamada K., Sax C.M., Norman B.,
Flanagan J.R., Ozato K., Westphal H., Platigorsky J.;
Regulation of the mouse alpha A-crystallin gene: isolation of a cDNA encoding a protein that binds to a cis sequence motif shared with the major histocompatibility complex class I gene and other genes.";
MOL. Cell. Biol. 10:3708-13708 (1990)
--- FUNCTION: TRANSCRIPTION FACTOR WHICH BINDS SPECIFICALLY TO THE PALINDROMIC SEQUENCE 5'-GGGAAATCCC-3' IN THE ALPHA-A CRYSTALLIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SUBCELLUIAR LOCATION: Nuclear.
SUBCELLUIAR LOCATION: Nuclear.
DOMAIN: CONTAINS TWO SETS OF 2 ZINC-FINGERS, WHICH ARE WIDELY
SEPARATED AND RECOGNIZE THE SAME DNA SEQUENCE. THERE IS A FIFTH
ZINC-FINGER IN-BETWEEN.
SIMILARITY: STRONG, TO HIVEP2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-0CT-1996 (Rel. 34, Created)
01-0CT-1996 (Rel. 34, Last sequence update)
10-0CT-2003 (Rel. 42, Last annotation update)
Zinc finger protein 40 (Transcription factor alphan-CRYBP1) (Alpha crystallin-binding protein I) (Alpha A-CRYBP1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE=95138112; PubMed=7836383; Brady J.P., Kantorow M., Sax C.M., Donovan D.M., Piatigorsky J.; Muxine transcription factor alpha A-crystallin binding protein I. Complete sequence, gene structure, expression, and functional inhibition via antisense RNA.; J., Biol. Chem. 270:1221-1229(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Mus musculus (Mouse).
Sukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia, Eutheria, Rodentia, Sciurognathi, Muridae, Murinae, Mus.
NCBI_TaxID=10090;
                                                                                                        1457 PI------GHQG----PI----GOEGAPGRPG--SPGLPGMP------
                                                                                                                                                                                                                                                                                                                   1597 AIAIAVHSQDVSIPHCPAGWRSLWIGYSFLMHTAAGDEGGGQSLV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2688 AA
                                                                                                                                                                                                                                                                                                                                                   LASGDRPPSVSSVHSEGDCN 2440
                                                                                                                                                                                                                                                                                                                                                                                     LEDFRATPFIECNGGRGTCH 1666
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE OF 2024-2688 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STANDARD:
                                                                                                                                                                                                                                                                                  2382 PITAADGRSD-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ZEP1 MOUSE
Q03172;
                                         1416
                                                                                                                                                                                                                                                                                                                                                   2421
                                                                                                                                                                                                                                                                                                                                                                                    1647
                                     a
                                                                       ò
                                                                                                    쉽
                                                                                                                                        à
                                                                                                                                                                          유
                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                  g
```

431GERRYPCVTCGPSTRTKSNLYKHKSHAHITKLGLVIOPBRAGGLFLSGECPKALSV 631 EAAPPPVVEDEBRESNVEBARALHASGNEVPRGECGGPA	DD 1385 PALVESQPSYSPSAVGGTAH	Δυτετικών του Δγ 1564 Ε	Db 1439 TIVLPSVNTVPFQAPRLPDMASA	Qy 1578 802		0y 1606 861	622 Db 1559	910 Qy 1653	Db 681		1004 Qy 1761	797 Db 1718	1061	848	1099	908 B06	1928	968 Db 1860	1163 Qy 1978	1028	1223 Qy		1275 QY 2056	1111 bb 2032	1329 2097 668	1171 Db 2091	1386	1220	1425 Qy 2187	1271 DB 2203 PSRS	1485	1324 Ov 2268	1506 - 2314 VQDSSVG	SSIAPSHVV 1384
		COS DAN DEBRINDACIVICANDE DENTE DA DEBRINDACIONES DE LA CONTRA DEL CONTRA DE LA CONTRA DE LA CONTRA DE LA CONTRA DE LA CONTRA DE LA CONTRA DE LA CONTRA DE LA CONTRA DE LA CONTRA DE LA CONTRA DE LA CONTRA DE LA CONTRA DE LA CONTRA DE LA CONTRA DE LA CONTRA DE LA CONTRA DE LA CONTRA DE LA CONTRA DE LA CONTRA DEL CONTRA DE LA CONTRA DE LA CONTRA DE LA CONTRA DE LA CONTRA DE LA CONTRA DE LA CONTRA DE LA CONTRA DE LA CONTRA DE LA CONTRA DE LA CONTRA DE LA CONTRA DE LA CONTRA DE LA CONTRA DE LA CONTRA DEL	487 HSDIEDSGES-DEEGLADGRONNPC/XDD.	DTESIPS-	:	803 TGAPTPPPAAPSPSAAPPV-VPKEEKEELTAAAPPVEEGEEGKPPAAEELAVD	577 LQVANPNPELPSPQSPRDLHVASILSHSASVSSLEMDESCHQK	862 VKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRA	623 IQSEGKPDSHSGTAHAQ-LQRQQATEDPQEQGGKLLLSPRSLGSTDSGYFSRSI	911 APODSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRANASPOK	967 KQRAAAIPPIQVTKVHBPPREDAAPTKPAPPAP	:: : : : : : : : : : : : : : : :	1005 LQPESDAPQQPGSSPRGKSRSPAPPADKEAFAAEAQKLPGDPPCWTSGLPPPVPPRE	798 TSSSSDIPKSP-FTPTEKSKQVFLLSVPSLDCLPITRSNSMPTTGYSAIPAN	1062 VIKASPHAPDPSAFSYAPPGHPLPLGLHDTARPVLPRP	:		909 GSGQSVDESCQGCPSSSEAGPVQSKAAQTPHLEKKKSHQGRGTWFECETCRNRYRKLENF	!	969 ENHKKFYCSELHGPKTKAAVREAEHGPAPGGAQPQVLHYRVAAPTAVWEQTPQIRKRRKM	1164 FSGVKQBQLSPRGQAGPPESLGVPTAQEASVLRGTALGSVPGGSITKGIPSTRVPSDSAI		TYRGSITHGT	 KHTSAT	1276 HVLSYEGGMSVTQCSKEDGRSSSGPPHETAAPKRTYDMME-GRVGRAISSASIEG	EVKRQGGGIS	1330 LMGRAIPPERHSPHHLKEQHHIRGSITQGIPRSYVEAQEDYLRREAKLLKREGTPPP	ÖddH		1221 EILVTEEPDRDLEAQSHDEEKSEKFTWPQRSETLSKLPFKLPPKKKRLRL	1426 HEIPREELRHTPELPLAPRPLKEGSITGGTPLKYDTGASTIGSKKHDVRSLIGSPGRIFP	1272 ARIEHSSTESSFESTLSRSLSRESSLSHAGSFSASLDLEDISKVEL	1486PVHPLDVMADARALERACYEE	1325 SKAEFLLIPLGSNTLSVPGSHREMRAASEQISCVPTLMEVSDFRSKSFDCGSIAPSHVV

	1385	PALVESQPSYSPSAVGGTAHVPLLER-RRGPLIRQISLNIASDSHLSPGSAAALQ 1438
	1564	EEGSLSSS 1577
	1439	TIVLPSVNTVPFQAPRLPDMASADCPAHTVHPQALAKDLQAEISSSSSTDTFPPQQLFGA 1498
	1578	KASQDRKLISTPREIAKSPHSTVPEHHP1605
	1499	HLLINKTNITSLSHQNTPLPLPVSAQGGKPDAPTACVSSTGEGSFAPKYQLQCQAFTSDQG 1558
	09	165
	1559	PGQAGADPCPASEAPPAKAADPMAKPCPLPPLELGLPRDEVLQKQLP
	1653	RHLAPNPTYPHLYPPYLI-RGYPDTAALBNRQTIINDYITSQQMHNTAT 1701
	1702	
•	1671	SMPKSQNYLQNASPTPEKELACKTVLPEVGQSVPVSESSPTVQKVSV 1717
	1761	LPTAPQPFSSRHSSSPLSPGGPTHLTKPTTTSSSERERDRDRERDRD 1807
	1718	GRLSPQQESSASSKRMLSPANSLDIAMEKHQKRAKDENGAVCSTNIRALELPSSRANESH 1777
	1808	REREKSILISTITVEHAPIWRPGTEQSSGSSGSSGGGGGSSSRPASHAHQHSPISPRT 1867
	1778	KOKKPVI.VRQLCTTEPLEGAALEQGACSASGRSSNKAANLTQVLP 1822
,. ,	1868	QDALQQRESVLHNTGMKGIITAVEPSKPTVLRSTSSPVRPAATFPPATHCPLGGTLDG 1927
	1823	TOSESSERVITATORG 1859
	1928 1860	VYPTLMEPV-LLPKEAPRVARPERPRADTGHAFLAKPPARSGLEPASSPSK 1977
••	1978	GSEPRPLVPRVSGHATIARTPAKNLAPHHASPDPPAPPA 2016
	1912	SDTQQPSFPSLKTATSFTWCYLLRQKALPLAQNDQKTSAYTGWTVSSSNPNPLGLPTKVA 1971
., .,	2017	SASDPHREKTQSKPFSIQELELRSLGYHGSSYSPEGVEP 2055 :: : : : : : : : :
••	2056	VSPVSSPSLTHDKGLPKHLEELDKSHLEGE
••	2032	
	2097	GBAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQRVVTLAQHISEVITQ 214
•	T602	GNLTKHMKSKAHSKKCVDLGVSVGLIDEQ 214
	2146	DYTRHHPQQLSAPLPAPLYSFPGASCPVLD
.,	2187	PPDHGAPARGSPHSEGGKRSPEPNKTSVLGGG 2218
••	2203	 VSSCFSGVHTDPMDILPRALLTKMTVLSTVQ
••	2219	LLYRD
	2263	:
, ''	2268	KLNTHNR
.,	2314	VQDSSVGLPPA-VAQLNPQPAARISSSVSPHPDSQDQKQQIILQPPPGLPSP 2364
	2310	GTEIFNMPAITGTGLMTYRSQAVQEHASTNMGLEAIIRKALMGKYDQWEESPPLSA 2365 : :

```
2461
                                                                                                                                                                          2424 GDRPPSVSSVHSEGDCNRRTPLTNRVWEDRPSSAGSTPFPY----NPLIMRLQ--AGV-M 2476
                                                                                                                                                                                                                     2522 SPAPPA----HIQGLQILNIALPTLIPSVGPVAVGTTGTPETTAPNSKAMELQMPAGQGH 2577
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A Nebliane Signature A Salve Bell, Adachi J., Bono H., Kondo S., R. Adachi Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S., R. Vikaido I., Osato N., Saito R., Suuki H., Yamanaka I., Kiyosawa H., A Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T., Ra Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J., Ra Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J., Ra Balke J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S., Bake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S., Ra Balke J.A., Fletcher C.F., Forrest A., Frezer K.S., RA Gasaeterland T., Gariboldi M., Gissi C., Godzik A., Gough J., RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D., RA Kamai H., Kawail H., Kawasawa Y., Kedzierski R.M., King B.L., RA Nagashima T., Numaca K., Okido T., Pavan M.J., Pertea G., Pesole G., RA Ravasi T., Numaca K., Okido T., Pavan W.J., Pertea G., Pesole G., Radelin A., Schneider C., Semple C.A., Setou M., Shimada K., Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K., Salutana R., Taylor M.S., Teasdale R., Wangner L., Taylor M.S., Teasdale R., Yang L., Yang L., Rangliawa M., Yang I., Yang L., Rang L., Yann Z., Zavolan M., Zimmer A., Carninci P., Hayatsu N.,
2366 NAFNPLNA---SASLPAAMPITAADGRSDHTLTSPGGG-----GKAKVSG-----
                                          STFVPIQAGPMOLTIPAVSVIHRTVGTSGDTITEASGSPNRPTGVAELSSVVPCIPIGQI
                                                                                                                              2462 HVPGLQNLSPPALQSLTSLGMETVNLVGLANATVGPQGHPPGLALNAVGLQVLANAPAQS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Zako M., Shinomura T., Ujita M., Ito K., Kimata K., "Expression of PG-M(V3), an alternatively spliced form of PG-M without a chondroitin sulfate attachment in region in mouse and human
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Mus musculus (Mouse).
Bukaryota, Metazoa, Chordata, Craniata, Vertebrata, Buteleostomi,
Mammalia, Butheria, Rodentia, Sciurognathi, Muridae, Murinae, Mus.
                                                                                                                                                                                                                                                                                                                                                                                                                                       062059; Q62058; Q9CUU0;
01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
15-MAR-2004 (Rel. 43, Last annotation update)
Versican core protein precursor (Large fibroblast proteoglycan)
(Chondroitin sulfate proteoglycan core protein 2) (PG-M).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A. (ISOFORMS V0; V1 AND V2).
STRAIN=C57BL/6, and Swiss Webster; TISSUE=Brain;
BEDLINE=9512251; PubMed=7822336;
Ito K., Shinomura T., Zako M., Ujita M., Kimata K.;
"Multiple forms of mouse Pd-M, a large chondroitin sulfate proteoglycan generated by alternative splicing.";
                                                                                                                                                                                                                                                                 2477 ASPPPPGLPAG-----SGPLAGPHHAWDEE 2501
                                                                                                                                                                                                                                                                                                             2578 SAEPPQGSPEGPQETPQTVSGPSA--DHARPED 2608
                                                                                                                                                                                                                                                                                                                                                                                                                                PRT; 3358 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              J. BIOL. Chem. 270:3914-3918 (1995).

[3]
SEQUENCE OF 1-1692 FROM N.A. (ISOFORM VI).

STRAIN=C57BL/6J; TISSUE=Skin;

MEDLINE=22354683; PubMed=12466851;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE=95181355; PubMed=7876137;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A. (ISOFORM V3).
                                                                                      2408 -RPSSRKAKSPA-----
                                                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAIN=C57BL/6;
                                                                                                                                                                                                                                                                                                                                                                                                                           PGCV MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                tissues.";
                                            2402
                                       셤
                                                                                                                            ద
                                                                                                                                                                                                                     g
                                                                                      ò
                                                                                                                                                                        à
                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                         셤
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                            ö
                                                                                                                                                                                                                                                                                     J. Biol. Chem. 274:20444-20449(1999).
-!- FUNCTION: May play a role in intercellular signaling and in connecting cells with the extracellular matrix. May take part in the regulation of cell motility, growth and differentiation. Binds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ISOId=0.4. Sequence=VSP 003087, VSP 003090;
-1 TISOIDE SPECIFICITY: V2 is found only in brain.
-1 DEVELOPMENTAL STAGE: Disappears after the cartilage development.
-1 DEVELOPMENTAL STAGE: Disappears after the cartilage development.
-1 SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
-1 SIMILARITY: Contains 2 link domains.
-1 SIMILARITY: Contains 1 E6F-like domains.
-1 SIMILARITY: Contains 1 Sushi (SCR) domain.
-1 SIMILARITY: Contains 1 Sushi (SCR) domain.
-1 SIMILARITY: Belongs to the aggrecan/versican proteoglycan family.
Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K., Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S., Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I., Wiyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A., Yasunishi A., Yoshino M., Waterston R., Lander B.S., Rogers J., Birney B., Hayashizaki Y.;
"Analysis of the mouse transcriptome based on functional annotation (60,770 full-length cDNAs.");
                                                                                                                                                                                                                                 Aspberg A., Adam S., Kostka G., Timpl R., Heinegaard D., "Fibulin-1 is a ligand for the C-type lectin domains of aggrecan and
                                                                                                                                                                                                                                                                                                                                                                  IsoId=Q62059-2; Sequence=VSP_003087, VSP_003088;
                                                                                                                                                                                                                                                                                                                                                                                                                                            Event=Alternative splicing, Named isoforms=4,
Comment=Additional isoforms seem to exist;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IsoId=062059-3; Sequence=VSP_003089;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IsoId=Q62059-1; Sequence=Displayed;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL, AKO14525; BAR99411.2; -...
HSSP; PO1132; 1EPG.
MGD, MGI.102899; GEPG2.
INTERPRO; IPRO0152; Asx hydroxyl_S.
INTERPRO; IPRO0162; Asx hydroxyl_S.
INTERPRO; IPRO01801; EGF_Z.
INTERPRO; IPRO01801; EGF_Z.
INTERPRO; IPRO01809; IGF_like.
INTERPRO; IPRO01809; IGF_like.
INTERPRO; IPRO01809; IGF_like.
INTERPRO; IPRO01809; IGF.
INTERPRO; IPRO01809; IGF.
INTERPRO; IPRO01809; IGF.
INTERPRO; IPRO01809; IGF.
INTERPRO; IPRO01809; IGF.
INTERPROS; IRRE; 2.
FEAN; PPO0009; EGF; 2.
FEAN; PPO0180; IGF; 2.
FEAN; PPO0180; IGF; 2.
FEAN; PPO0180; IGF; 2.
FEAN; PROO180; IGF; 2.
FEAN; PROD0180; INTERMODULE.
PRODOM; PD000918; Link; 2.
FRANT; SMO032; CCP; 1.
                                                                                                                                                                                                                 MEDLINE=99329059; PubMed=10400671;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    D28599; -; NOT ANNOTATED CDS.
D32040; BAA06802.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; D16263; BAA03796.1; -.
                                                                                                                                                                                               INTERACTION WITH FBLN1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Name=V1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Name=V3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Name=V0;
                                                                                                                                                                                                                                                                       versican.";
J. Biol. Ch
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Name=V2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL;
```

```
3358 AA; 366938 MW; 071B80026BC0762D CRC64;
                                                                                                                                                                                                                                             | : : | | : : | PGDEDNSTSKPLPSTEHAGFTKL---
                                                                                                                                                                                                                            207
  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     871
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1152
                                                                                                                                                                                                                                                      802
                                                                                                                                                                                                                                                                                                                                                                                                                                  435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  491
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    599
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    707
  S
                                                                                               a
                                                                                                                                                                                               В
                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                     8
                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       d
d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                           ò
                                                                                                                                                                         8
                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                             ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ⋩
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             .8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                   EGF-LIKE 1.
EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL)
                                                                                                                                                                                                                                                                GAG-ALPHA
(GLUCOSAMINOGLYCAN ATTACHMENT DOMAIN)
GAG-BETA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                             Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi; Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain; Hyaluronic acid; Alternative splicing.

1 20 POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /FTIGE-VSP 003090.
A -> G (IN REF. 3).
MISSING (IN REF. 3).
I -> T (IN REF. 3).
TVWNSNS -> QFGIQTA (IN REF. 3).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -> R (in isoform V1 and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Arital VSP 003089.
Missing (in isoform V3)
(FTIGEVSP 003000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (In isoform V2)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GECONAC.
GECONAC.
GELONAC.
GELONAC.
GELONAC.
GELONAC.
GELONAC.
GELONAC.
GELONAC.
GELONAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (In isoform V1)
                                                                                                                                                                                                               VERSICAN CORE PROTEIN. IG-LIKE V-TYPE.
                                                                                                                                                                                                                                                                                                                                                    BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              003088
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       003087
                                                                                                                                                                                                                                                                                                                           C-TYPE LECTIN SUSHI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FTIG=VSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FTIG=VSP
                                 SWART; SW00445; LINK; 2.

PROSITE; PS00010; ASX HYDROXYL; 1.

PROSITE; PS00010; ASX HYDROXYL; 1.

PROSITE; PS00041; C_TYPE_LECTIN 1; 1.

PROSITE; PS00022; EGF 1; 2.

PROSITE; PS01186; EGF 2; 1.

PROSITE; PS01186; EGF 3; 2.

PROSITE; PS01187; EGF 3; 2.

PROSITE; PS01187; EGF 3; 2.

PROSITE; PS01241; LINK; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Missing
SMART; SM00034; CLECT; 1.
SMART; SM00179; EGF_CA; 1.
SMART; SM00409; IG; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            126
348
1658
1680
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3143
3252
3254
3302
3315
3315
330
351
441
807
914
                                                                                                                                                                                                                                                                                                                                                                                                                  3067
3076
3087
3105
3114
3125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3052
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3052
                                                                                                                                                                                                                                                                                                                                                                                                                                3061
3078
3094
3099
3116
3132
3160
3228
3288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1679
2054
2244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3030
3332
3342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          126
348
1658
1674
                                                                                                                                                                                                                                                                                                                                                     DISULFID
DISULFID
DISULFID
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                         DISULFID
DISULFID
DISULFID
DISULFID
DISULFID
DISULFID
DISULFID
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CONFLICT
CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                   DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                      DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         JARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /ARSPLIC
                                                                                                                                                                                                               CHAIN
DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                     DOMAIN
```

J.

. 1209 EQRVFEEESSEKTTGDILPGESVTQHPVTTLIDIVAMKTESDIDHMTSKPPVTQPTRPSV 1268 GKSDQ-----DILEQ-- 1002 -----THLEATMSPGALRITGVSQGE------TQEEPQTPGSPFPTFSST 1041 374 950 91 YLPELGKSEMEFIESKRPRLELLPDPLLRPSPLLATGQPAGSEDLTKDRSLTGKLEPVSP 150 151 PSPPHTDPELELVPPRLSKEELIQNMDRVDREITMVEQQISKLKKKQQQLEEEAAK---- 206 801 267 HENIKINGAMRKKLILYFKRRNHARKQWKQKFCQRYDQLMEALEKKVERIENNPRRRAKE 326 -----EDLASGELTVTFHTSTSIGSAEKSASGEPTTGDRFLPTTSTEDQVINA 915 QEKETFREKFMOHPKNFGLIA----SFLERKTVAECVLYYYLTKKNENYKSLVRRSYRRR 490 551 DDTSGEDNDEKEAVASKG--RKTANSQGRRKGRITRSMANEANSEEAITP----- 598 652 ŤEVELLPY---SGDTTSVEGISTVÍY----PŚLQTDVTQGRERTETPRPELKKDPYTVD 703 375 IIDGLSEQENLEKQMRQLAVIPPMLYDADQQRIKFINMNGLMADPMKVYKDRQVMNMWSE 434 90 EGKK--ASLDYTDLGSG-LFEPRATELPKFPSTPSDISVFTAIDSLHRTPPLSPSSSFTE 704 EIPE-----SESSVERTES 742 VSPALTIEKLIGKPTEARDVEEMTTLTRLETDVTKSDKDVTRVHLTHSTLNVEVVTVSKW PPBPEKPVSPPPIESKHRSLVQIIYDENRKKAEAAHRILEGLGPQVELPLYNQPSDTRQY 830 -STIGING--KDKEIPSFTDGG-------GEYTLFPDGTPKPLEKVSE-----327 SKVREYYEKOPPEIRKOREL------OERMOSRVGORGSGLSMSAARSEHEVSE GKSQQQQQQQQQQQQQQQPMPRSSQEEKDEKEKEREREKEEKPEVENDKEDLLKEKT -QOSABLASMELNESSRWTEEEMETAKKGLLEHGRNWSAIARMVGSKTVSQCKNFYFNYK 658 KRONLDEILQOHKLKMEKERNARRKKKKAPAAASEEAAFPP-VVEDEEME--------ASGVSGNEEEMVEEAEALHASGNEVPRGECSGPATVNNSSDTESIPSPHTEAAKDTGQ 765 NGPKPPATLGADGPPPGPPTPPRRTS--RAPIEPTPASEATGAP----TPPPAPPSPSAP 819 PPVVPKEEKEEETAAAPPVEEGEEQKPPAA-----EELAVDTGKAEEPVKSECTEEA EEGPAKGKDAEAAEATAEGALK-----AEKKEGGSGRATTAK-SSGAPQDSDS--S ATCSAD----LSPRPSLLTPTGD 35 TDVGLLEYQHHSRDYASHLSPGSIIQPQRRRPSLLSEFQPGNERSQ----ELHLRPESHS Indels 1025; Gaps 1042 AVMAKETTAFEEGEGSTYTPSE------GRLMTGSERVPGLET----Length 3358; ---- PPVPL Query Match 2.3%; Score 310.5; DB 1; Best Local Similarity 17.8%; Pred. No. 0.005; Matches 516; Conservative 326; Mismatches 1024;

70	951 PRANASPQKPLDLKQLKQRAAAIPPIQVTKVHEPPREDAAPTKPAPPAPPAPPONN1.0 1006	ò	1858 HQHSPISPR1
đ	FILAE	qq	2265 LGTSILGDKN
ò		<i>\$</i>	1911 ATFPPATHCE
qq	TLSSHDGK	qq	2321 STLP
à		λō	1967 SGLEPASSPS
qq	SEGAIEOASOETILSRATEVAL	q _Q	2367 EAEEGATSPI
ζ		ò	1999 AKNLA
QQ	SWVEITPSOTVKF	අු	2427 ASSQGDSMLA
à		à	2051 EGVEPVSPVS
Ω		qq	2479 STLPSTEPYK
ζ		ò	2110 PES
Ор		ପ୍ଧ	2517 TESIIIDLDK
ò	1251 SPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETAAP 1307	à	2155 LSAPLPAPLY
g		qq	2563PVY
ò		λõ	2212 TSVLGGGEDG
g		qa	2608 ESFEGSGDAL
ò		ð	2254
q		q	2658 LHTFFPTAST
à	ASTT	ò	2294 LNTHNRNEPE
qo	FEPTGLTPSTVLDIEIAGVMSO	eg G	2716 LGHLEKTQEE
ò		ò	2348 KALMGKYDOWI
qq		dg	2776 PSDSTHYTF
ζ	1527 PVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSPVTWREPTPRLOEGSLS 1575	ò	2402 KAKVSGRPSSI
qq	::: VGMGGSDDERVRDTOTS	ପ୍ପ	2826VSTVTSSC
ò		ò	2436 EGDCNRRIPL1
ΩÞ	>	a a	2883 LLEELETSPTR
ò		à	2491 LAGPHHA
. <mark>д</mark>		qu	2935 LGGATQOPHSA
ò	1659 PTYPHLYPPYLIRGYPDTAALENROTIINDYITSQOMHHNTATAWAQRADWLRGLSP 1715	RESULT 100	T 100
qa	2059 LDFPQTWEPSKLWSKPEVNLDKQEIGRETVTKEKAQGQKTFESLHSSFAPEQTILETQSL 2118	NFH MOUSE	OUSE NFH MOUSE STAND
ò	1716 RESSLALNYAAGPRGIIDLSQVPHLPVLVPPTPGTPATAMDRLAYLPTA 1764		61959; 90 (Re
qu	2119 IÉTERQTSDYSMLTTLKTYITNKEVEEEGMSIAHMSTPGPGIKDLESYTTH 2169		01-AUG-1992 (Rel. 2 10-OCT-2003 (Rel. 4
λo	POPFSSRHSSSPLSPGGPTHLTKPTTTSSSBRBRDRBRBRDRERE 		Neurofilament triple (Neurofilament heavy
g	2170 PEAPGKSHSFSATALVTESGAARSVLMDSSTQBEE 2204		Mus musculus (Mouse)
8	1825 PIWRPGTEQSSGSSGSSGGGGGGSSRPASHSHA 1857	388	Eukaryota; Metazoa; Mammalia; Eutheria; NCBI TaxID=10090;
2	ZZUS SIKLIKŲKGVKLTNKESNADLSFSGLGSGGALPPLPTTSVNLTDMKQIISTLYAETSHMES 2264		[1]

```
PLGGTLDGVYPTLMEPVLLPKEAP----RVARPERPRADTGHAFLAKPPAR 1966
                                                                                                                                                                       SS------TIARTP 1998
TODALOORPSVLHNTGMKGIITAV-----EPSKPTVLRSTSTSSPVRPA 1910
                                                                                                                        ...----FVKLMDLERSPKQDPSGGKRKPKTHRPQTMSGLISNENSSAS 2366
                                                                                                                                                                                                            ;
PTAFLPQTYSVEMTKHFAPSESQPSDLFNVNSGEGSGEVDTLDLVYTSGTTQ 2426
                                                                                                                                                                                                                                                           APH---HASPDPPAPPASASDPHREKTQSKPFSIQELELRSLGYHGSSYSP 2050
                                                                                                                                                                                                                                                                                           SSPSLTHDKGLPKHLEELDKSHLEGE-LRPKQPGPVKLGGEAAHLPHLRPL 2109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -----QPSSSPLLQTAPGVKGHQRVVTLAQHISEVITQDYTRHHPQQ 2154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   YSF-PGASCPVLDLRRPPSDLYLPPPDHGAPARGSPHSE--GGKRSPEPNK 2211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GIEPVSPPEGMTEPGHSRSAVYPLLYRDGEQTE------- 2253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ----PSRMGSKSPGNTSQP-----PAFFSKLTESNSAMVKSKKQEINKK 2293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ||::::|| : : || TLHIPSKLTTASP-EIDKPNIEAISLDDIFESSTLSDGQAI-ADQSEVIST 2715
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EYNISQPGTEIFNMPAITGTGLMTYRSQAVQEHASTNM-----GLEAIIR 2347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | | : | | : | | : | | : | | EXERKXGGPSFQPEFFSGVGEVLTDPPAXVSIGSTYLIAQTLTELPNVVR 2775
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WEESPPLSANAFNPLNASASLPAA-MPITAADGRS-----DHTLTSPGGGG 2401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -EATPEVSSIA----ELSPQIPSSPFPVYVDNGVSKFPEVPHTSAQP---- 2825
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SRKA-----SVHS 2435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |:|:
SQKSIESPFKEVHANIEETIKPLGGNVHRTEPPSMSRDPALDVSEDESKHK 2882
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LTNRVWEDRPSSA----GSTPFPYNPLIMRLQAGVMASPPPPGLPAGSGP 2490
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           : | | : | : | | : | TKPETSQDFPNKAKDHIPGET-----VGMLAGIRTTESEPVITADDME 2934
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        IA 2497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3A 2945
```

MESULI 100
NFH MOUSE

TO NFH_MOUSE

TO NFH_MOUSE

TO P19246; 061959;
TO 1-NOV-1990 (Rel. 16, Created)

TO 1-NOV-1990 (Rel. 23, Last sequence update)

TO 1-AC-1992 (Rel. 24, Last annotation update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Neurofilament triplet H protein (200 kba neurofilament protein)

DE (Neurofilament heavy polypeptide) (NF-H).

GN NEFH OR NFH.

OS Mus muscratus (Mouse).

OC Bukaryota, Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;

OC Mammalia; Butheria; Rodentia; Sciurognathi; Muridae; Musrinae; Musr

SEQUENCE FROM N.A.

RP

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the BMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ibs.ch).
                                                                                                                                                                                       PTM: Phosphorylation seems to play a major role in the functioning of the larger neurofilament polypeptides (NF-M and NF-H), the levels of phosphorylation being altered developmentally and coincident with a change in the neurofilament function.

SIMILARITY: Belongs to the intermediate filament family.
CAUTION: Ref. 2 sequence differs from that shown in positions 534 to 716 and is shorter due to frameshifts.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 l coil; Neurone; Phosphorylation; Repeat.
HEAD.
                                                                             SEQUENCE. FROM N.A.
MEDLINE=89089138; PubMed=3145094;
Shneidman P.S., Carden M.J., Lees J.F., Lazzarini R.A.;
The structure of the largest murine neurofilament protein (NF-H) revealed by cDMA and genomic sequences.";
Brain Res. 464:217-231(1988).
           D., Grosveld
                              "Sequence and structure of the mouse gene coding for the largest neurofilament subunit.";
Gene 68:307-314(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LINKER 12.
COLL 2A.
LINKER 2.
COLL 2B.
K -> QA (IN REF. 2 AND 3).
A -> AR (IN REF. 2 AND 3).
S -> T (IN REF. 2 AND 3).
L -> G (IN REF. 2 AND 3).
P -> PREAKSP (IN REF. 3).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GLU-RICH (ACIDIC).
50 X 6 AA TANDEM REPEATS
GLU/LYS-RICH.
Julien J.-P., Cote F., Beaudet L., Sidky M., Flavell Mushynski W.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COIL 1A.
LINKER 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COIL 1B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    M23349; AAA39813.1; JOINED.
M24494; AAA39813.1; JOINED.
M24495; AAA39813.1; JOINED.
M35131; AAA39809.1; ALT_FRAME.
Z31012; CAA83229.1; -.
                                                                                                                                                                              STRAIN-Swiss Webster, TISSUE-Brain;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Coiled
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR001664; IF.
Pfam; PF00038; filament; 1.
PROSITE; PS00226; IF; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; M24496; AAA39813.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               filament;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    408
1087
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            517
886
1087
129
141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      239
283
283
283
133
133
189
281
551
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PIR; JT0368; QFMSH.
MGD; MGI:97309; Nefh.
                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Intermediate
                                                                      [2]
SEQUENCE. F
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  OMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SMBL;
```

K-S-P-A-E-A

```
AIPPIQVTKVHEPPREDAAPTKPAPPAPPPPQNLQPESDAPQQPGSSPRGKS----RSP- 1026
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SVQLHVPYSEHAKAPVGPVTM----GLPLPM------DPKKLAPFSGVKQEQLSPRG 1176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   625
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           680
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     914
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           971
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         --APPADKEAFAAEAQKLPGD---PPCWTSGLPFPVPPREVIKASPHAPDPSAFSYAPPG 1081
                                                                                                                                                                                                                   143
                                                                                                                                                                                                                                                     KORELOERMOSRVGQRGSGLSMSAAR----SEH---EVSEIIDGLSEQ----ENLEKQ 388
                                                                                                                                                                                                                                                                                                                                                  507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           567
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               288 FRVRLDRLSEAAKVNTDAMRSAQE----EITEYRRQLQARTT------ELEALKS- 332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        --TKESLERQRSELEDRHQADIASYQDAI-QQLDSELRNTKWEMAAQLREYQDLLNVKM 388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                389 AL---DIEIAAYRKILEGE---ECRIGFGPSPFSLTEGLPKIPSISTHIKVKSEEMIKVV 442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 735
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     502
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       736 CSGPATVNNSSDTESIPSPHTEAAKDTGONGPKPPATLGADGPPPGPPTPPRRTSRAPIE 795
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           547
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             796 PTPASEATGAPTPPPAPPSPSAPPPV-VPKEEKEETAAAPPVEEGEEQKPPAAEELAVD 854
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | : | |: | : | |: | | : | | SPAEAKSPAEAKSPAEAKSPAE-----VKSP------GEAKSPAEPKSPAEPKSPAEVKSPAE 703
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         880 КЕВVКSPVКВЕVКАКЕРРККVЕБЕКТІРТРКТЕАКЕSККDEAPKEAPKPKVEEKKETPTE 939
                                                                                                                                                                                 ARKQWKQ-----KFCQRYDQL--MEALEKKVERIENNPRRRAKESKVREYYEKQFPEIR
                                                                                                                                                                                                                                                                            --REVRE------MRGAVLRLGAARGQLRLEQEHLLEDIAHVRQRLDEEARQREEAEAAA
                                                                                                                                                                                                                                                                                                                                                                                                   KNFGLIASFLERKTVAECVLYYYLTKKNENYKSLVRRSYRR-RGKSQQQQQQQQQQQQQ
                                                                                                                                                                                                                                                                                                                                                                                                                                       EEVGELLGQIQGCGAAQAQAQ---AEARDALKCDVTSALREIRAQLEGHAVQSSLQSEEW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TGKAEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSSGAPQD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          389 MRQLAVIPPMLYDADQQRIKFINMNGLMADPMKVYKDRQVMNMWSEQEKETFREKFMQHP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QQQPMPRSSQEEKDEKEKEKEAEKEEEKPEVENDKEDLLKEKTDDTSGEDNDEKEAVASK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GRKTANSQGRRKGRI-TRSMANEANSEEAITPQQSAELASMELNESSRWTE-EEMETAKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GLLEHGRNWSAIARWVGSKTVSQCKNFY--FNYKKRQNLDEI--LQQH-KLKMEKERNAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RKKKK-----APAAASEEAAFPPWEDEEMEASGVSGNEEEMVEEAEALH-ASGNEVPRGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     443 EKSEKETVIVEGQTEEIRVTEGVTEEEDKEAQGQEGEEAEEGEEKEEEELAAATSPPAEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AASPEKETKSRVKEEAKSP--GEAKSPGE--AKSP----AEAKSPG-----EAKSPGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRANASPQKPLDLK---QLKQRAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AKSPAEVKSPGE-AKSPAAVKSPAEAKSPAAVKSPGEAKSPGEAKSPAEAKSPI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           763 EVKSPEKAKTPVKEGAKSPAEAKSPEKAKSPVKEDIKPPAEAKSPEKAKSPVKEGAKPPE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   823 KAKPL---DVKSPEAQTPVQEEATVPTDIRPPEQVKSPAKEKAKSPEKEEAKTSEKVAPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HPLPLGLHDTARPVLPRP-----PTISNPPPLISS-----AKHPSVLERQIGAISQGM
                                                                                                         2.3%; Score 309; DB 1; Length 1087;
                                                                                                                          Best Local Similarity 21.5%; Pred. No. 0.002;
Matches 245; Conservative 164; Mismatches 491; Indels
                                                                     MW; 57BAC76A38ED1CB9 CRC64;
MISSING (IN REF. 3).
G -> A (IN REF. 3).
V -> M (IN REF. 2 AND 3).
T -> N (IN REF. 2 AND 3).
                                                                         116612
 712
714
814
843
                                                                     1087 AA;
                                                                                                                                                                                                                                                     342
                                                                                                                                                                                 290
CONFLICT
CONFLICT
CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                         144
                                                                                                                                                                                                                                                                                                                                                                                                     449
                                                                                                                                                                                                                                                                                                                                                                                                                                         231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           508
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   568
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 681
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           503
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           915
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               658
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   972
                                                                                                         Match
                                                                                                           Query
                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    쉱
                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
```

1177 QAGPPESLGVPTAQEASVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPAD 1236		666666
	4-64	č
TO THE PROPERTY THE PROPERTY AND DESCRIPTION OF THE PROPERTY O		
1020 TKERKTTESBKBFE-KBKMONTANTANTANTANTANTANTANTANTANTANTANTANTAN	_	
		ć
_		
FOT THIRD TANDED TO THE PROPERTY OF THE PART		
		,
		5
		•
		ì
	•	Ē
	•	ì
		,
14.7 VIVV KINITUTION TO THE TENTE TO THE TRANSPORT TO THE		5
		ċ
		•
* * * * * * * * * *		
		2
		ĉ
>>34 3444 > 14 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4		
		7
		5
The state of the s		Ξ
		ć
-		

Search completed: April 15, 2004, 22:18: Job time : 201 secs